

STIC-Biotech/ChemLib

116229

From: Bunner, Bridget  
Sent: Monday, March 08, 2004 11:59 AM  
To: STIC-Biotech/ChemLib  
Subject: sequence search

Hi! I'd like to request a sequence search for case 09/555,296:

1. the amino acid sequence of SEQ ID NO: 4

Thanks!

Bridget Bunner

Art Unit 1647  
Rem 4C65  
(571) 272-0881  
mailbox 4C70

CRFE

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: \_\_\_\_\_  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 9, 2004, 13:32:25 ; Search time 23 Seconds  
(without alignments)  
469.123 Million cell updates/sec

Title: US-09-555-296B-4  
Perfect score: 1125  
Sequence: 1 MKQVLLITFVSALATQA.....DRKTYDIFNECYNGEPWL 209

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/6C.COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/6D.COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1125	100.0	209	US-09-618-919A-22	Sequence 22, Appl
2	290.5	25.8	190	US-09-618-919A-18	Sequence 18, Appl
3	265	23.6	200	US-09-618-919A-20	Sequence 20, Appl
4	252	22.4	190	US-09-618-919A-16	Sequence 16, Appl
5	102.5	9.1	284	US-09-134-000C-3594	Sequence 3594, Ap
6	91	8.1	202	US-07-807-022A-1	Sequence 1, Appl
7	84.5	7.5	497	US-09-107-532A-5279	Sequence 5279, Ap
8	83	7.4	510	US-09-134-000C-6635	Sequence 6635, Ap
9	83	7.4	2285	US-09-308-375-2	Sequence 2, Appl
10	81	7.2	659	US-08-894-818B-1	Sequence 1, Appl
11	81	7.2	659	US-08-894-818B-5	Sequence 5, Appl
12	81	7.2	659	US-09-445-472-12	Sequence 12, Appl
13	79.5	7.1	913	US-08-487-890A-5	Sequence 5, Appl
14	79.5	7.1	913	US-08-478-435-5	Sequence 5, Appl
15	79.5	7.1	913	US-08-337-483-5	Sequence 5, Appl
16	79.5	7.1	913	US-08-478-373-5	Sequence 5, Appl
17	79.5	7.1	913	US-08-474-671-5	Sequence 5, Appl
18	79.5	7.1	913	US-08-483-577A-5	Sequence 5, Appl
19	79.5	7.1	913	US-08-897-438-5	Sequence 5, Appl
20	79.5	7.1	913	US-08-637-654-5	Sequence 5, Appl
21	79.5	7.1	913	US-08-649-518-5	Sequence 5, Appl
22	78	6.9	1871	US-08-694-869-1	Sequence 1, Appl
23	78	6.9	1871	US-09-349-546-1	Sequence 1, Appl
24	78	6.9	1871	US-09-502-831-1	Sequence 1, Appl
25	77.5	6.9	652	US-09-489-038A-12461	Sequence 12461, A
26	77	6.8	640	US-09-328-352-8058	Sequence 8058, Ap
27	77	6.8	914	US-08-487-890A-11	Sequence 11, Appl

28	77	6.8	914	2	US-08-478-435-11	Sequence 11, Appl
29	77	6.8	914	2	US-08-337-483-11	Sequence 11, Appl
30	77	6.8	914	2	US-08-478-373-11	Sequence 11, Appl
31	77	6.8	914	3	US-08-474-671-11	Sequence 11, Appl
32	77	6.8	914	3	US-08-483-577A-11	Sequence 11, Appl
33	77	6.8	914	3	US-08-897-438-11	Sequence 11, Appl
34	77	6.8	914	4	US-08-637-654-11	Sequence 11, Appl
35	77	6.8	914	4	US-08-649-518-11	Sequence 11, Appl
36	76.5	6.8	406	1	US-08-434-881-2	Sequence 2, Appl
37	76.5	6.8	406	3	US-08-977-771-2	Sequence 2, Appl
38	76.5	6.8	406	3	US-09-361-773-2	Sequence 2, Appl
39	76	6.8	416	4	US-09-540-236-2388	Sequence 2388, Ap
40	75.5	6.7	391	4	US-09-107-532A-4419	Sequence 4419, Ap
41	75	6.7	179	4	US-09-328-352-6376	Sequence 6376, Ap
42	75	6.7	246	4	US-09-107-532A-5142	Sequence 5142, Ap
43	75	6.7	912	1	US-08-487-890A-9	Sequence 9, Appl
44	75	6.7	912	1	US-08-487-890A-7	Sequence 7, Appl
45	75	6.7	912	2	US-08-478-435-7	Sequence 7, Appl

## ALIGNMENTS

```

RESULT 1
US-09-618-919A-22
; Sequence 22, Application US/09618919A
; Patent No. 6617312
; GENERAL INFORMATION:
; APPLICANT: Pacesen, Guido Christiaan
; TITLE OF INVENTION: Vasoactive Amine Binding Molecules
; FILE REFERENCE: 2369-1-001CON
; CURRENT APPLICATION NUMBER: US/09/618,919A
; CURRENT FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: US 09/180,723
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: GB 9610484.0
; PRIOR FILING DATE: 1996-05-18
; PRIOR APPLICATION NUMBER: GB 9707844.8
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Dermacentor reticulatus
US-09-618-919A-22

Query Match      100.0%; Score 1125; DB 4; Length 209;
Best Local Similarity 100.0%; Pred. No. 2.8e-114;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MKQVLLITFVSALATQALETTSAKGPNLMAHEBLCKYDAMKSIDQVSVTVYLA 60
DB      1 MKQVLLITFVSALATQALETTSAKGPNLMAHEBLCKYDAMKSIDQVSVTVYLA 60
QY      1 KTYENDGSGSOFKLOVOETKREKEDTVSVFFRNASSPIKYNTVETVKAVFQY 120
DB      61 KTYENDGSGSOFKLOVOETKREKEDTVSVFFRNASSPIKYNTVETVKAVFQY 120
QY      121 GYKNIRNAIEYQVGGGINTDTLFTDGLCDVFFYNADQGCGLMWKSKHYKGVDPYCT 180
DB      121 GYKNIRNAIEYQVGGGINTDTLFTDGLCDVFFYNADQGCGLMWKSKHYKGVDPYCT 180
QY      181 FVENFCAKDKRTYDIFNEBCVYNGEPWL 209
DB      181 FVENFCAKDKRTYDIFNEBCVYNGEPWL 209

RESULT 2
US-09-618-919A-18
; Sequence 18, Application US/09618919A
; Patent No. 6617312

```

```

; GENERAL INFORMATION:
; APPLICANT: Paesen, Guido Christiaan
; APPLICANT: Nuttall, Patricia Anne
; TITLE OF INVENTION: Vasoactive Amine Binding Molecules
; FILE REFERENCE: 2369-1-001CON
; CURRENT APPLICATION NUMBER: US/09/618,919A
; CURRENT FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: US 09/180,733
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: GB 9610484.0
; PRIOR FILING DATE: 1996-05-18
; PRIOR APPLICATION NUMBER: GB 9707844.8
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Rhinipcephalus appendiculatus
US-09-618-919A-18

Query Match      25.8%; Score 290.5; DB 4; Length 190;
Best Local Similarity 36.0%; Pred. No. 1,3e-23;
Matches 72; Conservative 32; Mismatches 79; Indels 17; Gaps 9;

QY 7 LLLTFVSAALATQAEFTTSAKAGENPLMAHELLGKYQDAMKSIDQSVSTVYLAKTVEN 66
DB 3 LLLTSLALVLA-----LSQVKGNOFPMWADENANGAHQDMKSLKADVENVYIMVAKATYKN 57
QY 67 DTSGWSGQFKCLOVQEIERYEDYTVTSVFTF-RNASSPIKYNTETVYKAVFOYGYNI 125
DB 58 DP-VWGNDFTCVGMADVANDVEDKSIQAEFLFMNADTNWQF--ATEKYTAVMYGY-NR 113
QY 126 RNAIEYQVGGGINTDTLFTDELCDVFPVPMAD---QGCELMYKSHYKVPYCYFV 182
DB 114 ENARVETEDGQVFTDVIAYSD--DNCVVIYFPTDNGEGYEIMT--TDYDIPANCLNK 170
QY 183 FNVFCADKRTYDIFNEECV 202
DB 171 FNEY-AVGRTRDVFYSACL 189

RESULT 3
US-09-618-919A-20
; Sequence 20, Application US/09618919A
; Patent No. 6617312
; GENERAL INFORMATION:
; APPLICANT: Paesen, Guido Christiaan
; APPLICANT: Nuttall, Patricia Anne
; TITLE OF INVENTION: Vasoactive Amine Binding Molecules
; FILE REFERENCE: 2369-1-001CON
; CURRENT APPLICATION NUMBER: US/09/618,919A
; CURRENT FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: US 09/180,733
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: GB 9610484.0
; PRIOR FILING DATE: 1996-05-18
; PRIOR APPLICATION NUMBER: GB 9707844.8
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Rhinipcephalus appendiculatus
US-09-618-919A-20

Query Match      23.6%; Score 265; DB 4; Length 200;
Best Local Similarity 34.5%; Pred. No. 8,5e-21;
Matches 69; Conservative 34; Mismatches 71; Indels 26; Gaps 10;

QY 3 MCVVLLTFVSAALATQAEFTTSAKAGENPLMAHELLGKYQDAMKSIDQSVSTVYLAKT 62
DB 1 MCVVLLTFVSAALATQAEFTTSAKAGENPLMAHELLGKYQDAMKSIDQSVSTVYLAKT 62
```

```

DB 1 MKVLLIV--LGAALQONDA-----NPTANBAKLSYQDAMKSIQODONERYILAQA 51
QY 63 TYENDTSGWSGQFKCLOV--QEIERYEDYTVTSVFTFRNAS-SPIKYNTETVYKAVFO 119
DB 52 TOTTD-GVNGEFTCVSVTAERIKGKLNATI-----LYKNGLTDLKESHETTITVMKAYD 106
QY 120 YGKNIRNAIEYQVGGGINT--DTLFTDELCDVFPVPMAD--NADG-CELMYKSHYK 173
DB 107 Y---TTENIGKYEQTGTQTFEDVFPSDYKNCVILFPERKSDGDIHLMVSEDKID 163
QY 174 HVDYCTFVFNFCADKRT 193
DB 164 KIPDCKFTMAYFAQQQEXT 183

RESULT 4
US-09-618-919A-16
; Sequence 16, Application US/09618919A
; Patent No. 6617312
; GENERAL INFORMATION:
; APPLICANT: Paesen, Guido Christiaan
; APPLICANT: Nuttall, Patricia Anne
; TITLE OF INVENTION: Vasoactive Amine Binding Molecules
; FILE REFERENCE: 2369-1-001CON
; CURRENT APPLICATION NUMBER: US/09/618,919A
; CURRENT FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: US 09/180,733
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: GB 9610484.0
; PRIOR FILING DATE: 1996-05-18
; PRIOR APPLICATION NUMBER: GB 9707844.8
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Rhinipcephalus appendiculatus
US-09-618-919A-16

Query Match      22.4%; Score 252; DB 4; Length 190;
Best Local Similarity 33.5%; Pred. No. 2e-19;
Matches 68; Conservative 28; Mismatches 69; Indels 18; Gaps 8;

QY 3 MCVVLLTFVSAALATQAEFTTSAKAGENPLMAHELLGKYQDAMKSIDQSVSTVYLAKT 62
DB 1 MKVLLIVLALVLAISQVKA-----DKVWADENANGAHQDMKSLKADVENVYIMVAKATYKN 52
QY 63 TYENDTSGWSGQFKCLOVQEIERYEDYTVTSVFTFRNASPIKYNTETVYKAVFOYGY 122
DB 53 TYNDP-VWGNDFTCVGTAAQNLNDEKVENWEMFMNADTV-YGHTFEKATDPKMYGY 110
QY 123 KNIRNAIEYQVGGGINTDTLFTDELCDVFPVPMAD---QGCELMYKSHYKVPYCY 179
DB 111 -KRNATITQTEDEGVLIDVLAIFSD--DNCYVIALGPDGAGYELMA--TDYDIPASCC 166
QY 180 TFFVNFCAKDRKTYDIFNEECV 202
DB 167 LEKFNEMAA-GLPVRDVTSDCL 188

RESULT 5
US-09-134-000C-3594
; Sequence 3594, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
```

Tue Mar 9 14:18:19 2004

us-09-555-296b-4.rat

Page 3

PRIOR FILING DATE: 1997-08-15  
NUMBER OF SEQ ID NOS: 6812  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO: 3594  
LENGTH: 284  
TYPE: PRT  
ORGANISM: Enterococcus faecalis  
US-09-134-000C-3594

Query Match 9.1%; Score 102.5; DB 4; Length 284;  
Best Local Similarity 26.9%; Pred. No. 0.0066;  
Matches 60; Conservative 26; Mismatches 72; Indels 65; Gaps 15;

QY 16 LATQAEITSA-----KAGENPLMAHEELGKYQDAWMSI--DQGVSVTVYLAQTY-----64  
DB 68 LVGDYETGSSYGFAYVKKQNP---ELIKFFNAGLKVLKDNQ---TYDKILNNYLATGD 119  
QY 65 ENDTGSWGSQFCKLQVQEIIEKEDYTVTSVTFPRNASPIKYNN-----TEVYA 116  
DB 120 ENITQDAGEQMK-----KITPKKRYVIAADSTF---APPEFGNAQGDYVGIADVLYKR 170  
QY 117 V-----FQY-GYKIRNAIEY-QVGG---GLNTDTLIFTDGLCDVFPYFNADQG 162  
DB 171 AAELQGFVFEFFIFSSAVQAVESGQADGMVAMTITD-----DRKKAPDFSVFPFDG 225  
QY 163 CELWYKSH--YKHPDYCTFVNV-----FCADRKTYD 195  
DB 226 IGIAYKGNDRKISYDLDKGGKVGKIGTESADPLEKRNKKYD 268

#### RESULT 6

US-07-807-022A-1  
Sequence 1, Application US/07807022A  
Patent No. 531010

GENERAL INFORMATION:  
APPLICANT: Comolily, Thomas M.  
APPLICANT: Neepel, Michael  
TITLE OF INVENTION: Protein for Inhibiting Adhesion of  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESS: Merck & Co., Inc.  
STREET: P. O. Box 2000  
CITY: Rahway  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 08840

#### COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/807,022A  
FILING DATE: 19911210

#### ATTORNEY/AGENT INFORMATION:

NAME: Parr, Richard S.  
REGISTRATION NUMBER: 32,586  
REFERENCE/DOCKET NUMBER: 18143  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 594-4958  
TELEFAX: 908-594-4720  
TELEX: 138825

#### INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 202 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-07-807-022A-1

Query Match 8.1%; Score 91; DB 1; Length 202;  
Best Local Similarity 20.8%; Pred. No. 0.072;  
Matches 46; Conservative 33; Mismatches 82; Indels 60; Gaps 11;

QY 2 KMOVLLLFPSAALATQAEITSAKAGENPLMAHEELGKYQDAWMSI--DQGVSVTVYLAQ 61  
DB 1 KALILAVPISITYSVFTSDTISE-----EDLNKRNDRIMKTYGK--SKYFWLLR 49  
QY 62 TTYENDTGSWGSQFCKL-----QVQEI-----EKEDYTVTSYF--TFRNASPD 104  
DB 50 RTYAVD-----GKSLICRYGTVLKRDKNRHRIBQIMSGYDSQETTKIKYITTRGSG 105  
QY 105 IKYNNVETVAVRQYGYKIRNAIEYQVGGGLINTDTLIFTDGLCDVFPY-----156  
DB 106 DRNH-----MGVSLQGYNH--TGIEYK-----MYDDQGCALIKYTKDNRNP 147  
QY 157 PNADQCELMWYKSHYHPDY--CTFVNVFCADRKTYD 195  
DB 148 QNLKACEMWATANDANSVNSIAACEVYQRRCPNNSYVD 188

#### RESULT 7

US-09-107-532A-5279  
Sequence 5279, Application US/09107532A  
Patent No. 6583275

GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESS: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354

#### COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIOR APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997

#### ATTORNEY/AGENT INFORMATION:

NAME: Arimello, Pamela Deneka  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 5279:

#### SEQUENCE CHARACTERISTICS:

LENGTH: 497 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium

#### FEATURE:

NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...497  
SEQUENCE DESCRIPTION: SEQ ID NO: 5279:  
US-09-107-532A-5279

Tue Mar 9 14:18:19 2004

us-09-555-296b-4.rat

Page 4

```
Query Match      7.5%; Score 84.5; DB 4; Length 497;
Best Local Similarity 22.6%; Pred. No. 1.4;
Matches 47; Conservative 29; Mismatches 85; Indels 47; Gaps 9;

QY 26 KAGENPLMAHEELLGKYODAMKSIDQVSVTYVLAITYEND--TGSWGSQPKCQVQEI 83
DB 15 KPEGFIMGATPAANQCEGAMVDKGLSVSDYV---TDSPLPKKKTQDMHTHQOV 71
QY 84 ERKEDYTVTVTFRRNASSPIKYNVTEVKAVPFGYKRNRIEYVGGGINTDTL 143
DB 72 KAAQDP---NSKTYPRKHGNDPRYHFKEDIRLFAWGGKCYRMSIAW-----TR 118
QY 144 IF-----TDGELCDVFPVENADQ---GCELWKKSHY-----KAVPD 177
DB 119 IFPHGDETFNENAGLFYDQVFECLKYGIEPVVLSHYEMFLVITYEGWPNKLIQF 178
QY 178 YCTFVFNVPCKADKRT--YDIFNE-ECV 202
DB 179 YVRAETVFRKRYKRYKWTMFEINCV 206

RESULT 8
US-09-134-000C-6635
; Sequence 6635, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT FILING DATE: 1998-08-13
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6635
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; US-09-134-000C-6635

Query Match      7.4%; Score 83; DB 4; Length 510;
Best Local Similarity 20.2%; Pred. No. 2.1;
Matches 49; Conservative 42; Mismatches 77; Indels 74; Gaps 13;

QY 5 VLLITFVSAALATQATTSAKAGENPL--MAHE-ELGK--YODAMKSIDQVSVTVV 58
DB 9 VKITIASALKAIDKDRKSTSKRAETGIRYNNHENDIMNNRIIFYVD-----DEG-----V 58
QY 59 LAKTYENDGWSGQSKCQVQEI ERKEDYTVTVTFRRNASSPIK-----YNN----- 109
DB 59 LREKXASNVRIPIHGF-----PEYDQKTQYILSNPEYETENBEIKYALAEYNSFQ 113
QY 110 --VTETVKAVFQYGYK-----NIRNAIEYVGGGINTD-----TLIF 145
DB 114 VVLQELVEGSSQKFEFYVARTNAEDRLCFQVADSNVGVNENYELQRIORHYITTEI 173
QY 146 TDGELCDVFPVENADQCELMWKKSHYKHPDCTVFNVPFCADKRTYDINNEGCYVNG 205
DB 174 KDGETVDIHR-----AEVWTDQNVY-----FVADENKDYELDEAPEPI--NP 213
QY 206 EP 207
DB 214 RP 215

RESULT 9
US-09-308-375-2
; Sequence 2, Application US/09308375
; Patent No. 6300117
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
```

```

; TITLE OF INVENTION: Proteases From Gram-Positive Organisms
; FILE REFERENCE: GC394-PCT
; CURRENT APPLICATION NUMBER: US/09/308,375
; CURRENT FILING DATE: 1999-05-14
; EARLIER APPLICATION NUMBER: EP9719636.4
; EARLIER FILING DATE: 1997-09-15
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 2285
; TYPE: PRT
; ORGANISM: Bacillus subtilis
; US-09-308-375-2

Query Match      7.4%; Score 83; DB 4; Length 2285;
Best Local Similarity 23.3%; Pred. No. 19;
Matches 51; Conservative 27; Mismatches 79; Indels 62; Gaps 10;

QY 6 VLLITFVSAALATQATTSAKAGENPLMAHEELLGKYODAMKSID--QGVSVTYVLAITY 63
DB 743 LTVSTVYGAFALG-----WALSLLSSFAEAKKADDFEQSQQTVEALTT 790
QY 64 YENDTGSWGSQPKCQVQEI ERKEDYTVTVTFRRNASSPIKYNVTEVKAVFQ---Y 120
DB 791 NHDSTKLIQYK-----ELQVKSRSILTS-----DEQEVLYQTQLAQTFPPLVYK 838
QY 121 GY-----KNIRNAIEYVGGGINTDTLIFTDGLCDVFPVENADQCELMWYK 169
DB 839 GYDSQGNALIKTNKELEKAI-----NTXEVYALKQEFPSAKKTFEDASKR--IKK 889
QY 170 S-----HYKVPDYCTFVFNVPFCADKRTYDIFNEECYV 203
DB 890 SKDELKQYQIADVD-----KGRPKWLIADDDY 920

RESULT 10
US-08-894-818B-1
; Sequence 1, Application US/08894818B
; Patent No. 6261822
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: YAMAMOTO, Katsuhiko
; APPLICANT: MURTA, Masanori
; APPLICANT: ASADA, Kiyozo
; APPLICANT: TSUNASAWA, Shunmu
; APPLICANT: KATO, Ikumobin
; TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States of America
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,818B
; FILING DATE: 20-MAY-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03253
; FILING DATE: 07-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 33285/1995
; FILING DATE: 12-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
```

REGISTRATION NUMBER: 25,618  
REFERENCE/DOCKET NUMBER: TAKAKURA=1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 628-5197  
TELEFAX: (202) 737-3528  
INFORMATION FOR SEQ ID NO: 1  
SEQUENCE CHARACTERISTICS:  
LENGTH: 659 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-894-818B-1

Query Match 7.2%; Score 81; DB 3; Length 659;  
Best Local Similarity 25.9%; Pred. No. 5;  
Matches 53; Conservative 20; Mismatches 62; Indels 70; Gaps 13;

QY 41 KYODAMK-----SIDQVSV--TYVLAKTYENDTGSW--GSQFKCLQVOEIERKEEDYT 91  
DB 436 KYDDYAKLFTGVSADKGSATHTFDVSGATFVATLTYMDTSSDIDLILYDNGNEVDYS 495  
QY 92 VTSVFTFRNASSPIKTYNTE--TYKAVFOYGYKIRNAIEYQV-----GGG- 136  
DB 496 YTAAYGFEK---VGYNPTAGTWIVKVV---SYKGAAN---YQVDVVS DSGLSQSGGGN 545  
QY 137 -----LNTDTLIFTDGLCDVFPYRNADQGCGLMWKSKHYKHPDYCTFFVNFYC 187  
DB 546 PNPENPNPPTPTDTQFTT--GSVND--YMDTSD-----TFMTNANS 583  
QY 188 AKDRKTYDI-----FNEECVYNGEP 207  
DB 584 GATKITGDLFTDTSYNDLILYLYDP 608

RESULT 11  
US-08-894-818B-5  
Sequence 5, Application US/08894818B  
Patent No. 6261822  
GENERAL INFORMATION:  
APPLICANT: TAKAKURA, Hikaru  
APPLICANT: MORISHITA, Mio  
APPLICANT: YAMAMOTO, Katsuhiko  
APPLICANT: MITTA, Masenori  
APPLICANT: ASADA, Kiyozo  
APPLICANT: TSUNASAWA, Susumu  
APPLICANT: KATO, Ikunoshin  
TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Browdy and Neimark  
STREET: 419 Seventh Street N.W., Ste. 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: United States of America  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/894,818B  
FILING DATE: 20-MAY-1998  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03253  
FILING DATE: 07-NOV-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 323285/1995  
FILING DATE: 12-DEC-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Browdy, Roger L.

REGISTRATION NUMBER: 25,618  
REFERENCE/DOCKET NUMBER: TAKAKURA=1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 628-5197  
TELEFAX: (202) 737-3528  
INFORMATION FOR SEQ ID NO: 5  
SEQUENCE CHARACTERISTICS:  
LENGTH: 659 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-894-818B-5

Query Match 7.2%; Score 81; DB 3; Length 659;  
Best Local Similarity 25.9%; Pred. No. 5;  
Matches 53; Conservative 20; Mismatches 62; Indels 70; Gaps 13;

QY 41 KYODAMK-----SIDQVSV--TYVLAKTYENDTGSW--GSQFKCLQVOEIERKEEDYT 91  
DB 436 KYDDYAKLFTGVSADKGSATHTFDVSGATFVATLTYMDTSSDIDLILYDNGNEVDYS 495  
QY 92 VTSVFTFRNASSPIKTYNTE--TYKAVFOYGYKIRNAIEYQV-----GGG- 136  
DB 496 YTAAYGFEK---VGYNPTAGTWIVKVV---SYKGAAN---YQVDVVS DSGLSQSGGGN 545  
QY 137 -----LNTDTLIFTDGLCDVFPYRNADQGCGLMWKSKHYKHPDYCTFFVNFYC 187  
DB 546 PNPENPNPPTPTDTQFTT--GSVND--YMDTSD-----TFMTNANS 583  
QY 188 AKDRKTYDI-----FNEECVYNGEP 207  
DB 584 GATKITGDLFTDTSYNDLILYLYDP 608

RESULT 12  
US-09-445-472-12  
Sequence 12, Application US/09445472  
Patent No. 6358726  
GENERAL INFORMATION:  
APPLICANT: TAKAKURA, Hikaru  
APPLICANT: MORISHITA, Mio  
APPLICANT: SHIMOTO, Tomoko  
APPLICANT: ASADA, Kiyozo  
APPLICANT: KATO, Ikunoshin  
TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE  
FILE REFERENCE: TAKAKURA=6  
CURRENT APPLICATION NUMBER: US/09/445,472  
PRIOR FILING DATE: 1999-12-06  
CURRENT APPLICATION NUMBER: 151969/1997  
PRIOR APPLICATION NUMBER: 151969/1997  
FILING DATE: 1997-06-10  
NUMBER OF SEQ ID NOS: 33  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 12  
LENGTH: 659  
TYPE: PRT  
ORGANISM: Thermococcus celer  
US-09-445-472-12  
Query Match 7.2%; Score 81; DB 4; Length 659;  
Best Local Similarity 25.9%; Pred. No. 5;  
Matches 53; Conservative 20; Mismatches 62; Indels 70; Gaps 13;

QY 41 KYODAMK-----SIDQVSV--TYVLAKTYENDTGSW--GSQFKCLQVOEIERKEEDYT 91  
DB 436 KYDDYAKLFTGVSADKGSATHTFDVSGATFVATLTYMDTSSDIDLILYDNGNEVDYS 495  
QY 92 VTSVFTFRNASSPIKTYNTE--TYKAVFOYGYKIRNAIEYQV-----GGG- 136  
DB 496 YTAAYGFEK---VGYNPTAGTWIVKVV---SYKGAAN---YQVDVVS DSGLSQSGGGN 545  
QY 137 -----LNTDTLIFTDGLCDVFPYRNADQGCGLMWKSKHYKHPDYCTFFVNFYC 187

Tue Mar 9 14:18:19 2004

us-09-555-296b-4.rat

Page 6

Db 546 ENNPNPNTPTTDTOTFT-GSVND--YWDJSD-----TTTMMVNS 583  
QY 188 AKDKRTYDI-----FNEECVNGEP 207  
Db 584 GATKITGDLTFDTSYNDLIDLILYDP 608

RESULT 13

US-08-487-890A-5  
Sequence 5, Application US/08487890A  
Patent No. 5708149  
GENERAL INFORMATION:  
APPLICANT: Loommore, Sheena  
APPLICANT: Harkness, Robin  
APPLICANT: Schryvers, Anthony  
APPLICANT: Chong, Pele  
APPLICANT: Gray-Owen, Scott  
APPLICANT: Yang, Yan-Ping  
APPLICANT: Mardin, Andrew  
APPLICANT: Klein, Michel  
TITLE OF INVENTION: Transferrin Receptor Genes  
NUMBER OF SEQUENCES: 147  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: 6th Floor, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,890A  
FILING DATE: 07-JUN-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/175,116  
FILING DATE: 29-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,968  
FILING DATE: 08-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-466 MIS:jb  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 913 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-487-890A-5

Query Match 7.1%; Score 79.5; DB 1; Length 913;  
Best Local Similarity 21.9%; Pred. No. 12;  
Matches 48; Conservative 32; Mismatches 60; Indels 79; Gaps 13;

QY 11 FVSALATQAEITSAKAGENPLMAHEILGKYD-AWKS-----IDGVSVTVYLAATY 64  
Db 593 YVDLIGLRMYDVSRKANESTI-----SVGKRNFSWNTGIVIKPTWMLDLSYRLS-TGF 646  
QY 65 EN-----DTGSWSQFKLCQVOEIERKE-----DYTVSVTFPRNASS 103  
Db 647 RNPSPAFMYGMYRGKDTDYVIGKFK-----PETSRRNQEFGLAKDQFGNIEISHPSNA-- 700  
QY 104 PIKRYNTEIVYKAVFQ-----YCKRIKRAIBYGVGGGANTIDTLIFDDELCDVYF 155

Db 701 ---YKNLAFAEELSKNCTTGKNGYGHANMAKLTV-----CVNTAQDPRNG----- 745  
QY 156 VENADQCELMYKSHYKRPD--YCTFERNVCAKDRK 192  
Db 746 -----LW-----KRIPYGWATFAYNRVAVKQOK 769

RESULT 14

US-08-478-435-5  
Sequence 5, Application US/08478435  
Patent No. 5922323  
GENERAL INFORMATION:  
APPLICANT: Loommore, Sheena  
APPLICANT: Harkness, Robin  
APPLICANT: Schryvers, Anthony  
APPLICANT: Chong, Pele  
APPLICANT: Gray-Owen, Scott  
APPLICANT: Yang, Yan-Ping  
APPLICANT: Mardin, Andrew  
APPLICANT: Klein, Michel  
TITLE OF INVENTION: Transferrin Receptor Genes  
NUMBER OF SEQUENCES: 147  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sim & McBurney  
STREET: Suite 701, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/478,435  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/337,483  
FILING DATE: 08-NOV-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/175,116  
FILING DATE: 29-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-462 MIS:VG  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 913 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-478-435-5

Query Match 7.1%; Score 79.5; DB 2; Length 913;  
Best Local Similarity 21.9%; Pred. No. 12;  
Matches 48; Conservative 32; Mismatches 60; Indels 79; Gaps 13;

QY 11 FVSALATQAEITSAKAGENPLMAHEILGKYD-AWKS-----IDGVSVTVYLAATY 64  
Db 593 YVDLIGLRMYDVSRKANESTI-----SVGKRNFSWNTGIVIKPTWMLDLSYRLS-TGF 646  
QY 65 EN-----DTGSWSQFKLCQVOEIERKE-----DYTVSVTFPRNASS 103

Db 647 RNPFAENYGRVGGKDTDYIGKFK---PETSROEFGALAKDFGNIISHFSNA-- 700  
QY 104 PIKYNNVETVKAVFQ-----YGYKINRNALIEYOVGGGLITPTLJFTDGLCDVRY 155  
Db 701 ---YRNLIAPABEISKNGTGTGKNGYGHNAQNALV---GVNITQOLDENG----- 745  
QY 156 VPNDQGCGLMWKSHKHVPD--YCTFVFNVCARDK 192  
Db 746 -----LM-----KRIPGWIATFAINRVKVDQK 769

QY 156 VPNDQGCGLMWKSHKHVPD--YCTFVFNVCARDK 192  
Db 746 -----LM-----KRIPGWIATFAINRVKVDQK 769  
Search completed: March 9, 2004, 13:36:04  
Job time : 25 secs

RESULT 15  
US-08-337-483-5  
; Sequence 5, Application US/08337483  
; Patent No. 5922562  
; GENERAL INFORMATION:  
; APPLICANT: Loommore, Sheena  
; APPLICANT: Harkness, Robin  
; APPLICANT: Schryvers, Anthony  
; APPLICANT: Chong, Pele  
; APPLICANT: Gray-Owen, Scott  
; APPLICANT: Yang, Yan-Ping  
; APPLICANT: Murdin, Andrew  
; APPLICANT: Klein, Michel  
; TITLE OF INVENTION: Transferrin Receptor Genes  
; NUMBER OF SEQUENCES: 147  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Sim & McBurney  
; STREET: Suite 701, 330 University Avenue  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5G 1R7  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/337,483  
; FILING DATE: 08-NOV-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stewart, Michael I  
; REGISTRATION NUMBER: 24,973  
; REFERENCE/DOCKET NUMBER: 1038-410 MIS:jb  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (416) 595-1155  
; TELEFAX: (416) 595-1163  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 913 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-337-483-5

Query Match 7.1%; Score 79.5; DB 2; Length 913;  
Best Local Similarity 21.9%; Pred. No. 12;  
Matches 48; Conservative 32; Mismatches 60; Indels 79; Gaps 13;

QY 11 FVSALATQAERTSKAGENPLMAHEILGKYOD-AWKS-----IDQVSVTVVLAKTV 64  
Db 593 YVDLGLGMRIVDSRIKANESTI-----SVGKFKNPSMTGIVIKPTWLDSTRLS-TGF 646  
QY 65 EN-----DTGSWGSQFKCLQVQELERKE-----DYTVTSVPTFRNASS 103  
Db 647 RNPFAENYGRVGGKDTDYIGKFK---PETSROEFGALAKDFGNIISHFSNA-- 700  
QY 104 PIKYNNVETVKAVFQ-----YGYKINRNALIEYOVGGGLITPTLJFTDGLCDVRY 155  
Db 701 ---YRNLIAPABEISKNGTGTGKNGYGHNAQNALV---GVNITQOLDENG----- 745



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 9, 2004, 13:33:35 ; Search time 33 Seconds

(without alignments)  
1337,304 Million cell updates/sec

Title: US-09-555-296b-4

Perfect score: 1125  
Sequence: 1 MKQVLLILTFVSAALATQA.....DRKTYDIFNECVNGEPMWL 209

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 21153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubppaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubppaa/PCUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubppaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubppaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubppaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/1/pubppaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubppaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubppaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	282	25.1	171	US-10-085-572-7	Sequence 7, Appl1
2	282	25.1	171	US-10-087-195-7	Sequence 7, Appl1
3	251.5	22.4	182	US-10-085-572-8	Sequence 8, Appl1
4	251.5	22.4	182	US-10-087-195-8	Sequence 8, Appl1
5	239	21.2	172	US-10-085-572-6	Sequence 6, Appl1
6	239	21.2	172	US-10-087-195-6	Sequence 6, Appl1
7	102.5	9.1	722	US-09-815-242-10796	Sequence 10796, A
8	97.5	8.7	221	US-09-728-914-4	Sequence 4, Appl1
9	84	7.5	217	US-08-728-914-22	Sequence 22, Appl1
10	83	7.4	2285	US-09-932-183A-2	Sequence 2, Appl1
11	82	7.3	381	US-10-369-493-5666	Sequence 5666, Ap
12	81.5	7.2	627	US-10-389-566-2357	Sequence 2357, Ap
13	81	7.2	659	US-10-090-624-12	Sequence 12, Appl1
14	80.5	7.2	738	US-10-291-583-83	Sequence 83, Appl1
15	79.5	7.1	738	US-10-291-583-92	Sequence 92, Appl1

16	79.5	7.1	738	US-10-291-583-95	Sequence 95, Appl1
17	79.5	7.1	738	US-10-423-704A-2	Sequence 2, Appl1
18	79.5	7.1	913	US-10-043-344-5	Sequence 5, Appl1
19	78.5	7.0	733	US-10-291-583-87	Sequence 87, Appl1
20	78.5	7.0	738	US-10-291-583-80	Sequence 80, Appl1
21	77.5	6.9	506	US-10-369-493-5064	Sequence 5064, Appl
22	77.5	6.9	685	US-10-291-583-113	Sequence 113, App
23	77.5	6.9	733	US-10-291-583-86	Sequence 86, Appl1
24	77.5	6.9	733	US-10-291-583-90	Sequence 90, Appl1
25	77.5	6.9	738	US-10-291-583-79	Sequence 79, Appl1
26	77.5	6.9	738	US-10-291-583-81	Sequence 81, Appl1
27	77.5	6.9	738	US-10-291-583-84	Sequence 84, Appl1
28	77.5	6.9	738	US-10-291-583-85	Sequence 85, Appl1
29	77.5	6.9	738	US-10-291-583-91	Sequence 91, Appl1
30	77.5	6.9	738	US-10-291-583-93	Sequence 93, Appl1
31	77.5	6.9	738	US-10-291-583-94	Sequence 94, Appl1
32	77	6.8	631	US-10-043-344-11	Sequence 7764, Ap
33	77	6.8	914	US-10-043-344-11	Sequence 11, Appl1
34	76.5	6.8	413	US-10-380-077-18	Sequence 18, Appl1
35	76.5	6.8	575	US-10-104-047-3822	Sequence 3822, Ap
36	76.5	6.8	738	US-10-291-583-82	Sequence 82, Appl1
37	76.5	6.8	1778	US-10-238-075-749	Sequence 749, App
38	76	6.8	403	US-10-369-493-11803	Sequence 11803, A
39	75	6.7	518	US-09-769-734-56	Sequence 56, Appl1
40	75	6.7	736	US-10-291-583-99	Sequence 99, Appl1
41	75	6.7	912	US-10-043-344-7	Sequence 7, Appl1
42	75	6.7	912	US-10-043-344-9	Sequence 9, Appl1
43	74.5	6.6	381	US-10-369-493-21237	Sequence 21237, A
44	74.5	6.6	389	US-10-094-944-22	Sequence 22, Appl1
45	74.5	6.6	492	US-10-369-493-23526	Sequence 23526, A

## ALIGNMENTS

RESULT 1				
US-10-085-572-7				
Publication No. US20020151499A1				
GENERAL INFORMATION:				
APPLICANT: Nuttall, Patricia, Ann				
TITLE OF INVENTION: Treatment of Conjunctivitis				
FILE REFERENCE: 2488-1-003				
CURRENT APPLICATION NUMBER: US/10/085,572				
CURRENT FILING DATE: 2002-02-27				
PRIOR APPLICATION NUMBER: PCT/GB00/03282				
PRIOR FILING DATE: 2000-08-24				
PRIOR APPLICATION NUMBER: 9920674.0				
NUMBER OF SEQ ID NOS: 8				
SOFTWARE: FastSeq for Windows Version 4.0				
SEQ ID NO 7				
LENGTH: 171				
TYPE: PRT				
ORGANISM: FS-HBP 2				
US-10-085-572-7				
Query Match				
Best Local Similarity 37.5%; Pred. No. 1.4e-22;				
Matches 66; Conservative 29; Mismatches 69; Indels 12; Gaps 8;				
QY				
31 PVAHEHLLGKYDANKSIDGVSVTVYLAKTYYENDTSGWSQFKCLOVOIERKEDY 90				
Db				
3 PDWAEKANAQADDAKSKLADVENYVVKATYKNDP-VMGADFTCVGVANVDYNEEK 61				
QY				
91 TVTSVVF--ENASPSKYTNVETVAKVQYGYKNIIRNAIEYOVGGGLNITTLTFTDSE 149				
Db				
62 SIQKEFFPANNADITNMQF--ATBKTVAKMGTG-NRENAFRLETEDGQVFTVIAXSD-D 117				
QY				
150 LCDVFYVFNAD--OGCELVVKKSHYKAVPDVCTFVNFVCAKDRKTYDIFNEECV 202				
Db				
118 NCDVLYVFGTDGNEGVELWT--TDYDNPANCLNKFNKY-AVGRSTRDVFVTSACL 170				

```
RESULT 2
US-10-087-195-7
; Sequence 7, Application US/10087195
; Publication No. US20020193306A1
; GENERAL INFORMATION:
; APPLICANT: Nuttall, Patricia, Anne
; TITLE OF INVENTION: Treatment of Allergic Rhinitis
; FILE REFERENCE: 2488-1-004
; CURRENT APPLICATION NUMBER: US/10/087,195
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: PCT/GB00/03287
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 9920673.2
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 171
; TYPE: PRT
; ORGANISM: FS-HBP 2
US-10-087-195-7

Query Match      25.1%; Score 282; DB 13; Length 171;
Best Local Similarity 37.5%; Pred. No. 1.4e-22;
Matches 66; Conservative 29; Mismatches 69; Indels 12; Gaps 8;

QY 31 PMAHEELLGKYODAWKSIDQGVSVTVYLAATTYENDTGSWGSQFCKLOVGEIRKSEEDY 90
DB 3 PMAHEELLGKYODAWKSIDQGVSVTVYLAATTYENDTGSWGSQFCKLOVGEIRKSEEDY 90
QY 91 TTVTSVFTF-RMASSPIKYNVTETKAVFOYGYKNIRNAIEYOVGGGLNTTDLTIFDGE 149
DB 62 SIOAEPLFMNNADMTNQF--ATEKTAIVVMYGY-NRENAFRETEDQGVFDVIAYSD-D 117
QY 150 LQGVFYPNAD---GQCELMVKKSHYKHPDYCTFVFNVCXKDKRTYDINSECV 202
DB 118 NCDVITVPGTGNBEGIELMT--TDYDINPANCINKRNEY-AVGRERIDVITSACL 170

RESULT 3
US-10-085-572-8
; Sequence 8, Application US/10085572
; Publication No. US20020151499A1
; GENERAL INFORMATION:
; APPLICANT: Nuttall, Patricia, Ann
; APPLICANT: Paesen, Guido, Christiaan
; TITLE OF INVENTION: Treatment of Conjunctivitis
; FILE REFERENCE: 2488-1-003
; CURRENT APPLICATION NUMBER: US/10/085,572
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: PCT/GB00/03282
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 9920674.0
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 182
; TYPE: PRT
; ORGANISM: MS-HBP 1
US-10-085-572-8

Query Match      22.4%; Score 251.5; DB 13; Length 182;
Best Local Similarity 35.3%; Pred. No. 3.3e-19;
Matches 61; Conservative 29; Mismatches 66; Indels 17; Gaps 8;

QY 30 NPLMAHEELLGKYODAWKSIDQGVSVTVYLAATTYENDTGSWGSQFCKLOV--OEIRKE 87
DB 1 NPTMAHEAKLGSYQDAMKSIQDQNKRYVLAQATQTTD-GVWGEFFTCVSVTAERIKGKK 59
```

```
QY 88 EDYTVTSVTFRNAS-SPIKYNTETKAVFOYGYKNIRNAIEYOVGGGLNTT--DTLI 144
DB 60 LNATI---LYKNHGLTDLKHSHETITVMAAYD---TTENGKYEFGRTQTFEDVFEV 112
QY 145 FTDSGLCVFYPV---NADQG-CELMVKKSHYKHPDYCTFVFNVCXKDKRT 193
DB 113 FSDYKNCVITVPERGSDGDEYELWSEDKIDKIPDCKFTMAVFAQOQEKI 165

RESULT 4
US-10-087-195-8
; Sequence 8, Application US/10087195
; Publication No. US20020193306A1
; GENERAL INFORMATION:
; APPLICANT: Nuttall, Patricia, Anne
; APPLICANT: Paesen, Guido, Christiaan
; TITLE OF INVENTION: Treatment of Allergic Rhinitis
; FILE REFERENCE: 2488-1-004
; CURRENT APPLICATION NUMBER: US/10/087,195
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: PCT/GB00/03287
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 9920673.2
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 182
; TYPE: PRT
; ORGANISM: MS-HBP 1
US-10-087-195-8

Query Match      22.4%; Score 251.5; DB 13; Length 182;
Best Local Similarity 35.3%; Pred. No. 3.3e-19;
Matches 61; Conservative 29; Mismatches 66; Indels 17; Gaps 8;

QY 30 NPLMAHEELLGKYODAWKSIDQGVSVTVYLAATTYENDTGSWGSQFCKLOV--OEIRKE 87
DB 1 NPTMAHEAKLGSYQDAMKSIQDQNKRYVLAQATQTTD-GVWGEFFTCVSVTAERIKGKK 59
QY 88 EDYTVTSVTFRNAS-SPIKYNTETKAVFOYGYKNIRNAIEYOVGGGLNTT--DTLI 144
DB 60 LNATI---LYKNHGLTDLKHSHETITVMAAYD---TTENGKYEFGRTQTFEDVFEV 112
QY 145 FTDSGLCVFYPV---NADQG-CELMVKKSHYKHPDYCTFVFNVCXKDKRT 193
DB 113 FSDYKNCVITVPERGSDGDEYELWSEDKIDKIPDCKFTMAVFAQOQEKI 165

RESULT 5
US-10-085-572-6
; Sequence 6, Application US/10085572
; Publication No. US20020151499A1
; GENERAL INFORMATION:
; APPLICANT: Nuttall, Patricia, Ann
; APPLICANT: Paesen, Guido, Christiaan
; TITLE OF INVENTION: Treatment of Conjunctivitis
; FILE REFERENCE: 2488-1-003
; CURRENT APPLICATION NUMBER: US/10/085,572
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: PCT/GB00/03282
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 9920674.0
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 172
; TYPE: PRT
; ORGANISM: FS-HBP1
US-10-085-572-6

Query Match      21.2%; Score 239; DB 13; Length 172;
```

```
Best Local Similarity 34.5%; Pred. No. 7.2e-18;
Matches 61; Conservative 24; Mismatches 82; Indels 10; Gaps 7;

QY 29 ENPLMAHEELIGKQDAMKSIDGVSTYVLAITYENDGSGQFKLQVGEIERKEE 88
DB 1 DKPWADDEAANGEHQDAMKHLQKLVSENYDLIKATYKNDP-VGNDPFTCVGTAAQNLNED 59
QY 89 DYTTSVFTFNNASSPIKYNYVTEVKAVFQYGYKNIRNAIEYQVGGGLNTDTLLFTDG 148
DB 60 EKAVFAWFMFNNDLV-YQHTFEKATPDKNYGY-NKENAITVQTEDEGVLTVLAFSD- 116
QY 149 ELCDVFYFNAD---QGCFLMWKSHYKHPDYCTFVFNFCAKDKRTVDIFNEECV 202
DB 117 DNCYVYALGPDGSGAGAYELMA--TDYTDVPASCLKEFENYAA-GLPVRDVIYSDCL 170

RESULT 6
US-10-087-195-6
Sequence 6, Application US/10087195
Publication No. US20020193306X1
GENERAL INFORMATION:
APPLICANT: Nuttall, Patricia, Anne
APPLICANT: Paesen, Guido, Christiaan
TITLE OF INVENTION: Treatment of Allergic Rhinitis
FILE REFERENCE: 2488-1-004
CURRENT APPLICATION NUMBER: US/10/087,195
CURRENT FILING DATE: 2002-03-01
PRIOR APPLICATION NUMBER: PCT/GB00/03287
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: 9920673.2
PRIOR FILING DATE: 1999-09-01
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 172
TYPE: PRT
ORGANISM: FS-HBPI
US-10-087-195-6

Query Match 21.2%; Score 239; DB 13; Length 172;
Best Local Similarity 34.5%; Pred. No. 7.2e-18;
Matches 61; Conservative 24; Mismatches 82; Indels 10; Gaps 7;

QY 29 ENPLMAHEELIGKQDAMKSIDGVSTYVLAITYENDGSGQFKLQVGEIERKEE 88
DB 1 DKPWADDEAANGEHQDAMKHLQKLVSENYDLIKATYKNDP-VGNDPFTCVGTAAQNLNED 59
QY 89 DYTTSVFTFNNASSPIKYNYVTEVKAVFQYGYKNIRNAIEYQVGGGLNTDTLLFTDG 148
DB 60 EKAVFAWFMFNNDLV-YQHTFEKATPDKNYGY-NKENAITVQTEDEGVLTVLAFSD- 116
QY 149 ELCDVFYFNAD---QGCFLMWKSHYKHPDYCTFVFNFCAKDKRTVDIFNEECV 202
DB 117 DNCYVYALGPDGSGAGAYELMA--TDYTDVPASCLKEFENYAA-GLPVRDVIYSDCL 170

RESULT 7
US-09-815-242-10796
Sequence 10796, Application US/09815242
Patent No. US2002061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trewick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.01A
CURRENT APPLICATION NUMBER: US/09/815,242
```

```
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 1410
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10796
LENGTH: 722
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-815-242-10796

Query Match 9.1%; Score 102.5; DB 9; Length 722;
Best Local Similarity 26.9%; Pred. No. 0.042;
Matches 60; Conservative 26; Mismatches 72; Indels 65; Gaps 15;

QY 16 LATQAEFTSA-----KAGENPLMAHEELIGKQDAMKSI-DGVSTYVLAITY-----TETVA 64
DB 204 LVGDKETGSSYGYPAVKKGNDP-----ELIKFVAGLKNKNDG--TYDKILNNVYATGD 255
QY 65 ENDGSGWQFKLQVGEIERKEEDYVTSVFTFNNASSPIKYNY-----TETVA 116
DB 256 ETTNODAGEQMK-----KITPKKEKTVIASDSTF---APFERQNAQSGYVGIDVLYR 306
QY 117 V-----FOY-GYKNIRNAIEY-QVGG--GLNTDTLLFTDSELCDFVFNADG 162
DB 307 AAEQGFTEFEKFIGFSSAVQAVESGQADGVAVGNTITD-----DRKKAFFDSVPYFDG 361
QY 163 CELMWKSH--YKAVPDYCTFVFNV-----FOCKDKRTVD 195
DB 362 IQIAVKKGNKIKSYDDLKGGKVGKIGTESADFEKKNKKID 404

RESULT 8
US-09-728-914-4
Sequence 4, Application US/09728914
Patent No. US20010046499A1
GENERAL INFORMATION:
APPLICANT: KANTOR, FRED S.
APPLICANT: FIRSIG, EROL
APPLICANT: DAS, SUBRATA
TITLE OF INVENTION: TICK ANTIGENS AND COMPOSITIONS AND METHODS COMPRISING
FILE REFERENCE: Y0-107
CURRENT APPLICATION NUMBER: US/09/728,914
CURRENT FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: 60/169,048
PRIOR FILING DATE: 1999-12-03
PRIOR APPLICATION NUMBER: 60/240,716
PRIOR FILING DATE: 2000-10-16
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 221
TYPE: PRT
ORGANISM: Ixodes scapularis
US-09-728-914-4

Query Match 8.7%; Score 97.5; DB 9; Length 221;
Best Local Similarity 23.4%; Pred. No. 0.028;
Matches 49; Conservative 43; Mismatches 80; Indels 37; Gaps 14;
```

**Tue Mar 9 14:18:20 2004**

us-09-555-296b-4.rapb

Page 4

```

QY 10 TFSALANLOAE-----TSKAGENLMAHEELG--QDAMKSDGVSATVLA 60
Db 31 TYSGTGTTTPTGTRARMTVTPAPEED-----SKYEONATRYE--MATOMK 80
QY 61 KTYENDTGSMQSFCKLOVIEERKEEDYIVTSV-FTFRNASSPIKTYNTET--VKAV 117
Db 81 MRYTD-VTPDSGNPVQCNFRWMEK-----TPTNSFQYRYSKN-SMETIDETLLKDI 134
QY 118 FQYGYKNIRAIYQVGGGLNTDPLL-FTDBELCDYVYVPAAG---CELMYK-SHY 172
Db 135 GEHGFPYVNMFTPTPG--IATDNLVATSNVNTCTVLRIPPTNGERHCJDMANNTLS 191
QY 173 KAVPDYCTFENVFCAKDRKTYDIFNEEC 201
Db 192 QETPDDCNKFFEXY-NTQIYKRVYVPSK 219

```

```

RESULT 9
US-09-728-914-22
; Sequence 22, Application US/09728914
; Patent No. US20010046499A1
; GENERAL INFORMATION:
; APPLICANT: KANTOR, FRED S.
; APPLICANT: PIKRIQ, EROL
; APPLICANT: DAS, SUBRATA
; TITLE OF INVENTION: TICK ANTIGENS AND COMPOSITIONS AND METHODS COMPRISING
; TITLE OF INVENTION: THEM
; FILE REFERENCE: YU-107
; CURRENT APPLICATION NUMBER: US/09/728,914
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/169,048
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: 60/240,716
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 217
; TYPE: PRF
; ORGANISM: Ixodes scapularis
; US-09-728-914-22

```

```

Query Match Similarity 7.5%; Score 84; DB 9; Length 217;
Best Local Similarity 24.1%; Pred. No. 0.82;
Matches 27; Conservative 17; Mismatches 48; Indels 20; Gaps 3

QY 99 RNASSGIRKNTYETKAVFQVYGVNINALEYGVGINITDILFETQS----- 148
DB 79 RNSS-----YEFKRLTSMISVGEQSPSHTEVEITGTSKSKSVKTKGYEYEDNTY 134

QY 149 --ELCDVFYVPMND---OGCELMYKSKHRYKVPDYCTFPVANNFCADKRTY 194
DB 135 ADTRCLILIRISRTQVFLRSCLLAWKCKFLKLPALHCKCFELDFVCKMRKRD 186

```

```

RESULT 10
US-09-932-183A-2
; Sequence 2, Application US/09932183A
; Patent No. US20020127641A1
; GENERAL INFORMATION:
; APPLICANT: Estell, David A.
; TITLE OF INVENTION: Proteases From Gram-Positive Organisms
; FILE REFERENCE: GC394CI-US
; CURRENT APPLICATION NUMBER: US/09/932.183A
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/308,375
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/US98/18828
; PRIOR FILING DATE: 1998-09-08
; PRIOR APPLICATION NUMBER: EP9719636.4
; PRIOR FILING DATE: 1997-09-15
; NUMBER OF SEQ. ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0

```

```

? SEQ ID NO 2
? LENGTH: 2285
? TYPE: PRT
? ORGANISM: Bacillus subtilis
US-09-932-193A-2

Query Match          7.4%   Score 83; DB 9; Length 2285;
Best Local Similarity 23.3%   Pred. No. 26;
Matches 51, Conservative 27, Mismatches 79, Indels 62, Gaps 10.

```

```

QY      6 VILLIFVSAALTAQAEITTSAGAEINFLTAHEHLLIGKQDAMKSID--QGSVTVYVAKIT 63
Dh      743 ILVSTLVGAPALALG-----WALESLISFAEAKKADDPGEGQQGNVEALIT 750

QY      64 YENDTGSWGSGPKCLQVQIEKREEDVTVISVTFNNAASPIKTYNTFVKAFAVQ---Y 120
Dh      791 NKDSTDKLIQYK-----ELQKVESRSILTS-----DEBEQELVQTQOLACFPALVYK 838

QY      121 GR-----KNIRNAIEYQVGGGLNITDILLTFDGEICDVFYVPMADQGGELMVKK 159
Dh      839 GVDGSGNALIKTNKELEKAIK-----NIXEYIALKKQEFSDSAKTFEDASKK--IKK 889

QY      170 S-----HYKGVDPDYCTFVFNVFCACARDKTYDIFNEECYV 203
Dh      890 SKDELKQYKQIDVND-----KGRFKMLIDDDDDY 920

```

```

RESULT 11
US-10-369-493-5686
Sequence 5686, Application US/10369493
Publication No. US2003023675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 5686
LENGTH: 381
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-10-369-493-5686

```

Query Match	7.3%	Score 82;	DB 15;	Length 381;
Best Local Similarity	21.8%	Pred. No. 3;		
Matches	29;	Mismatches	56;	Indels 28;
				Gaps 3;

```

QY      9  LTVSMAALTOAETTSAKGENPTLAHAEELG-----KYQAMKSIDGVSVTY  57
DB      35  LEIVDPQTKAASAEINNESEVENELADBPMDQDDNRPVIFPLHHIDIMNFKYKAVASF  94
QY      58  VLAQTTIYENDUTSGWSQFCLOVQIIEKEDYVTSYFERNASSPT-----KXVNV  110
DB      95  TVEEVLDLQKMDM-----EKMGDEQYFISRIALFAASDGIENMLCERSNE  144
QY      111  TETKAVFOYGYK  123
DB      145  VQVSEARFFYGFO  157

```

RESULT 12  
US-10-389-566-2357  
; Sequence 2357, Application US/10389566  
; Publication No. US2004002520A1  
; GENERAL INFORMATION:

Query Marc36 7.2% Score 81, DB 13, Length 659;  
 Best Local Similarity 25.9%; Pred. No. 8.2;  
 Matches 53; Conservative 20; Mismatches 62; Indels 70; Gaps 13

QY 41 KYDAMK-----SIDQGVSV-TYVLANTYENDTSGM--GSGFKCLQVQLERKEEDYT 91  
 Db 436 KYDDVAKLFTFGSVADKGSATHTFPVSATFTVTLTYMDTGSSDIDLTYDPNGNEVDSS 495  
 QY 92 VTSVFETFRNASSPIKTYNTE--TVKRVFQYGYKRNARNAIEYOV-----GGG- 136  
 Db 496 YTYLYGFER---VGIYNFTAGTIVYKV---STKGAN---YQVDVDSGSLSSGGGAN 545  
 QY 137 -----LNTDTLLFTDGLCDVFYPVNAQSGCELWKKSHYKRVDPDYCTFVFNVFC 187

RESULT 15  
US-10-291--583-92  
Sequence 92, Application US/10291583  
Publication No. US20030138772A1  
GENERAL INFORMATION:  
APPLICANT: Gao, Guangping  
APPLICANT: Wilson, James M.  
APPLICANT: Alvira, Mauricio  
TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus  
TITLE OF INVENTION: Sequences and Isolating No. US20030138772A1 Sequences Ident  
FILE REFERENCE: UPN-02735USA  
CURRENT APPLICATION NUMBER: US/10/291,583  
CURRENT FILING DATE: 2002-11-12  
PRIOR APPLICATION NUMBER: US 60/350,607  
PRIOR FILING DATE: 2001-11-13  
PRIOR APPLICATION NUMBER: US 60/341,117  
PRIOR FILING DATE: 2001-12-17  
PRIOR APPLICATION NUMBER: US 60/377,066  
PRIOR FILING DATE: 2002-05-01  
PRIOR APPLICATION NUMBER: US 60/386,675

```
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 92
; LENGTH: 738
; TYPE: PRT
; ORGANISM: capsid protein of AAV serotype, clone 43.1
US-10-291-583-92

Query Match
  7.1%; Score 79.5; DB 14; Length 738;
  Best Local Similarity 28.3%; Pred. No. 14;
  Matches 45; Conservative 19; Mismatches 60; Indels 35; Gaps 12;

QY 13 SAALATQAEITTSKAGENPLMAHEBLLGKYQD-AMKSIDQGVSVTVYLAKTYYENDTGSW 71
DB 233 STWLGDRVITTTSTRT-----WA-----LPTVNNHLVKQISNGTSGST-NDNTYFGYSTPW 282
QY 72 G-----SQFKLQVQELERKEDYTVTSVFTFENASSPIKYYNVTETVKAVPQ-YGYKNIR 126
DB 283 GYFDENRFHC---HFSRDMQRLINNMGEFRPKRLNPKLFNI--QYKEVTQNEGTXTIA 336
QY 127 NAIYYQVGGSLNITDL-ITDGLCDVFTVP-NADQGC 163
DB 337 N-----NLITSTIQVFTDSEY-QLPYVPGSAHQGC 364
```

Search completed: March 9, 2004, 13:36:50  
Job time : 34 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: March 9, 2004, 13:30:31 ; Search time 20 Seconds

(without alignments)  
1005.201 Million cell updates/secTitle: US-09-555-296B-4  
Perfect score: 1125  
Sequence: 1 MKMGVLLLTFFVSALATQA.....DRKTYDIFNECVNGEPWL 209Scoring table: BLOSUM62  
Gap: 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: PIR\_78:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	* Query Match	Length	DB	ID	Description
1	96.5	8.6	621	2	F97862	heat shock protein
2	96	8.5	650	2	G70169	heat shock protein
3	87	7.7	639	2	F71358	probable heat shock
4	86	7.6	2747	2	B49132	fat facets (faf) s
5	85	7.6	656	2	A87112	heat shock protein
6	84.5	7.5	621	2	H71645	heat shock protein
7	84.5	7.5	917	2	AF1394	autolysin, amidase
8	83	7.4	2285	2	T12796	probable transglyc
9	82.5	7.3	318	2	C69192	stomatin-like prot
10	82	7.3	163	2	C97308	probable acetyltra
11	82	7.3	360	2	T17996	hypothetical prote
12	82	7.3	381	2	T18876	hypothetical prote
13	82	7.3	390	2	D82922	conserved hypotet
14	82	7.3	637	2	T08530	trig protein - Ent
15	82	7.3	637	2	S22992	trig protein - Ent
16	81.5	7.2	627	2	D71729	dnax-type molecula
17	80.5	7.2	621	2	A71961	chaperone and heat
18	79.5	7.1	621	2	B64546	hypothetical prote
19	79.5	7.1	725	2	AG2547	hypothetical prote
20	78.5	7.0	455	2	A75121	hypothetical prote
21	78.5	7.0	864	2	S60431	hypothetical prote
22	78	6.9	215	2	A26669	nodulin-22 prote
23	78	6.9	523	1	O4CKA3	cytochrome P450 52
24	78	6.9	1871	2	S27938	hypothetical prote
25	77.5	6.9	223	2	T12076	ribonuclease (EC 3
26	77.5	6.9	506	2	T25870	hypothetical prote
27	77	6.8	241	2	T26909	hypothetical prote
28	77	6.8	331	2	D81409	probable periplasm
29	77	6.8	780	2	H84453	probable heat shock

30	77	6.8	914	2	S70906	transferrin-bindin
31	76.5	6.8	587	2	E87420	poly-beta-hydroxyb
32	76.5	6.8	635	2	T46407	probable RNA helic
33	76.5	6.8	807	2	T18454	hypothetical prote
34	76.5	6.8	2819	2	A90551	conserved hypotet
35	76	6.8	250	2	A55122	deoxycytidine kina
36	76	6.8	510	2	A71516	hypothetical prote
37	76	6.8	510	2	A64706	lipase-like prote
38	76	6.8	531	2	T18741	hypothetical prote
39	76	6.8	598	2	AG0890	probable arylsulfa
40	76	6.8	781	2	S49340	heat-shock protein
41	76	6.8	1031	2	T38411	probable GTPase ac
42	76	6.8	1066	2	B95037	hyaluronidase [imp
43	76	6.8	1068	2	S73091	hypothetical prote
44	76	6.8	1277	2	B84517	hypothetical prote
45	75.5	6.7	820	2	D81382	probable transmemb

## ALIGNMENTS

## RESULT 1

F97862 heat shock protein hspg [imported] - Rickettsia conorii (strain Malish 7)

C/Species: Rickettsia conorii

C/Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 22-Oct-2001

C/Accession: F97862

R/Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.

Science 293, 2093-2098, 2001

A/Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.

A/Reference number: A97700; MUID:21442074; PMID:11557893

A/Accession: F97862

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-621 &lt;KUR&gt;

A/Cross-references: GB:AE006914; PTDN:ALM03840.1; PID:915620441; GSPDB:GN00173

C/Genetics:

A/Gene: hspg

C/Superfamily: heat shock protein 90

Query Match 8.6%; Score 96.5; DB 2; Length 621;  
Best local similarity 21.0%; Pred. No. 1.2; Indels 35; Gaps 10;  
Matches 43; Conservative 46; Mismatches 81;

QY	11	FVSALATQAEITTSARAGNP--LMAHEELGKY--QDAWKSIDGVSVTVYLAATYEN 66
DB	127	YSSFVYADKVTYSRAGEDKHIW-ESDGLGEYTVSDSKFTGTETVILIKK---EE 182
QY	67	DTGSMGSGPFCLOVQIEKEBEDYVTSYTF-----RNAASPIKYNVETKAV 117
DB	183	DT--FDDHFR---LKHIVSYSDHIAVPIYFDEAGNNEIQUNSAALTRKSEITTEQ 237
QY	118	FOYGYKNIRNAIE-----YGVGGELNITDILFTDGLCDVFPVPAADQCELMVKK- 169
DB	238	YKPFYSLSYAIDDPMTIHNNKNGAIBETNLIFFPSSKTFPLFH-PKRRKVKYIKRV 296
QY	170	----SHKAVPDYCTFVNVFCAKD 190
DB	297	FISDENIDILPSYLRFLRGVDSED 321

## RESULT 2

G70169 heat shock protein 90 (hspg) homolog - Lyme disease spirochete

N/Alternate names: C62.5 heat shock protein

C/Species: Borrelia burgdorferi (Lyme disease spirochete)

C/Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 20-Aug-1999

C/Accession: G70169; I40247

R/Fraser, C.M.; Casjens, S.; Huang, W.M.; Suton, G.G.; Clayton, R.; Lathigra, R.; W

son, D.; Peterson, J.; Keriavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; V

Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B

Nature 390, 380-386, 1997

A/Authors: Smith, H.O.; Venter, J.C.

A/Title: Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*.  
A/Reference number: A70100; MUID:98065943; PMID:9403685  
A/Accession: G70169  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-650 <RLE>  
A/Cross-references: GB:AE001157; GB:AE000783; NID:g2688471; PID:AG66919.1; PID:g268847  
A/Experimental source: strain B31  
R/Otani, C.; Davidson, B.E.; Saint Girons, I.; Old, I.G.  
Microbiology 140, 2931-2940, 1994  
A/Title: Conservation of gene arrangement and an unusual organization of rRNA genes in  
A/Reference number: 140241; MUID:95111614; PMID:7812434  
A/Accession: 140247  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 531-609 <RES>  
A/Cross-references: GB:U32145; NID:g476004; PIDN:AC41403.1; PID:g476594  
C/Genetics:  
A/Genes: htpG  
C/Superfamily: heat shock protein 90  
Query Match 8.5%; Score 96; DB 2; Length 650;  
Best Local Similarity 21.5%; Pred. No. 1.4;  
Matches 46; Conservative 34; Mismatches 92; Indels 42; Gaps 8;  
Cy 11 FVSALATQ-AETTSAKAGBNP-LMAHELLIGYQDAWKSIDGVSVTVLAKTYEND 67  
Db 155 FVSAPFVSEKVEVTSKKALESADAYIWSGDKGYBIEKAKESGTEKLYNKGLE-- 212  
Cy 68 TGSWGSQFKLOY-----QEIERKEEDYTVTSVTFPRNA 101  
Db 213 ---YANKMKIGELIKKYSNHNPIYIKSEPIMDGQKQGEIEBEKLNATTALMTKN- 268  
Cy 102 SSPIKTYNVTETVKAFFQYGYKNIRNAIEYQGGGINTDTLITFDGELCDVFFVFNADQ 161  
Db 269 KSEIKAEVNEFYKNT-TPDYENPLMHHTKAGNLEVTNLFYVPSKAPDLVY-FMTKP 326  
Cy 162 GCELMVKK-----SHYKVPDYCTFVFNVCAND 190  
Db 327 GVKLFINKRITITDSGSLPNTLRFKGIIDCOD 360  
RESULT 3  
F71258  
probable heat shock protein 90 (htpg) - syphilis spirochete  
C/Species: *Treponema pallidum* subsp. *pallidum* (syphilis spirochete)  
C/Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 05-Nov-1999  
C/Accession: F71258  
R/Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin  
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uetereback, T.; Mcd  
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
Science 281, 375-388, 1998  
A/Title: Complete genome sequence of *Treponema pallidum*, the syphilis spirochete.  
A/Reference number: A71250; MUID:98352770; PMID:9665876  
A/Accession: F71258  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-639 <COL>  
A/Cross-references: GB:AE001265; GB:AE000520; NID:g3323295; PIDN:AC65938.1; PID:g332330  
A/Experimental source: strain Nichols  
C/Genetics:  
A/Genes: TP0984  
C/Superfamily: heat shock protein 90  
Query Match 7.7%; Score 87; DB 2; Length 639;  
Best Local Similarity 22.9%; Pred. No. 9;  
Matches 50; Conservative 26; Mismatches 52; Indels 90; Gaps 10;  
Cy 5 VVLLTFVSALATQAEITTSAKAGBNPLMAHELLIGKYD-----AKSIDGVSV 54  
Db 182 VVLLTSGENSEEPATR-----WRLSEVIRKYSDDHAFPIYLHYLQKEYDXDGA 228  
Cy 55 VTVYVAKTYENDTGSWGSQFKLOYQEIERKEEDY-----TVTSVTFPRNASPIKTYN 109

Db 229 VTDQKQKQDVVNDAGALMKRPS-----ELKEDYHRFQYTLT-----KDSFTPLLYVH 277  
Cy 110 VTEYVKAFFQYGYKNIRNAIEYQGGGINTDTLITFDG--ELCDVFFVP-----NA 159  
Db 278 TK-----AEQGEVYTLFVPAKAPFDLFA 303  
Cy 160 D--QCELMVKK-----SHYKVPDYCTFVFNVCAND 190  
Db 304 DYKPEVKLFVKRVFTTDEKELLPVYLRFAVGVIDSD 341  
RESULT 4  
B49132  
fat facets (fat) splice form 1 - fruit fly (*Drosophila melanogaster*)  
C/Species: *Drosophila melanogaster*  
C/Date: 19-Dec-1993 #sequence\_revision 25-Apr-1997 #text\_change 01-Dec-2000  
C/Accession: B49132; A49132  
R/Fischer-Vize, J.A.; Rubin, G.M.; Lehmann, R.  
Development 116, 985-1000, 1992  
A/Title: The fat facets gene is required for *Drosophila* eye and embryo development.  
A/Reference number: A49132; MUID:93202020; PMID:1295747  
A/Contents: isogenic at  
A/Accession: B49132  
A/Status: preliminary  
A/Molecule type: nucleic acid  
A/Residues: 1-2747 <FIS>  
A/Cross-references: GB:U04959; NID:g157411; PIDN:AAF01345.1; PID:g6013474  
A/Note: sequence inconsistent with the nucleotide translation  
A/Note: sequence extracted from NCBI backbone (NCBIN:127836, NCBIN:129008, NCBI:P:12  
A/Accession: A49132  
A/Status: preliminary  
A/Molecule type: nucleic acid  
A/Residues: 1-2704, 'VT', 2707, 'ANNV', <FIS>  
A/Cross-references: GB:U04958; NID:g157410; PIDN:AAF01346.1; PID:g6013475  
A/Note: sequence extracted from NCBI backbone (NCBIN:127836, NCBIN:129008, NCBI:P:12  
C/Keywords: alternative splicing  
Query Match 7.6%; Score 86; DB 2; Length 2747;  
Best Local Similarity 22.9%; Pred. No. 75;  
Matches 54; Conservative 28; Mismatches 72; Indels 82; Gaps 12;  
Cy 6 VVLLTFVSALATQAEITTSAKAGBNPLMAHELLIG-----KYDA-----WKSIDG 52  
Db 1756 VVLLKHVQAFIHALGSHALQYVPRGLMTHFKLGEFVNLEQDDAVEFWSLIESIDEG 1815  
Cy 53 VSVTVYVAKTYENDT--GSWGSQFKLOYQEIERKEEDYTVTSVTFPRNASPIKTYN 110  
Db 1816 LK--ALQGPQLMNTLGGSFSDQKICQECPRYSKBPVSFVSV-DIRNHS-----L 1865  
Cy 111 TETVAVAFQYGYKNIRNAIEYQGGGINTDTLITFDGEL--CDVEYVFNADGCELMV 167  
Db 1866 TSLDQ-----YVNGELLEADAVHCDKCDKV-VTV 1896  
Cy 168 KSHYKVP-----DY--CTFVFNVCADKDKTYTDFNECYVNGSPW 208  
Db 1897 KRCVVKQDPPVLAQLKRFEDYERVCAIKN-----DYFEPRIIDDEPY 1942  
RESULT 5  
A87112  
heat shock protein Hsp90 family [imported] - *Mycobacterium leprae*  
C/Species: *Mycobacterium leprae*  
C/Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001  
C/Accession: A87112  
R/Cole, S.T.; Eglmeier, K.; Pakhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.  
R.; Davies, R.M.; Devlin, K.; Dutchoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Hol  
eam, M.A.; Rutherford, K.M.  
Nature 409, 1007-1011, 2001  
A/Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R  
A/Title: Massive gene decay in the leprosy bacillus.  
A/Reference number: A86909; MUID:21128732; PMID:11234002  
A/Accession: A87112



A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-656 <STO>  
A/Cross-references: GB:AL450380; NID:g13093407; PIDN:CA030574.1; GSPDB:GN00147  
A/Genetics:  
A/Gene: htpg  
A/Superfamily: heat shock protein 90

Query Match 7.6%; Score 85; DB 2; Length 656;  
Best Local Similarity 21.3%; Pred. No. 14;  
Matches 49; Conservative 38; Mismatches 79; Indels 64; Gaps 9;

QY 11 FVSALATQAEITTSKAGENPL--WAHELLGKYQDAWKSID---QGVSVTYVLAATYVE 65  
DB 137 YSSFWANKEVLTTRKAGETATRWSD---GEATYTESYDEAPQGTSTVTLHKPEDEE 193  
QY 66 NDTGWSGQFKLCQV-----QETERR-----BEDTVTSVTFEN 100  
DB 194 DELHDYSEWKIRELVKYSDFIAMPTRMEVERAPATSDGEGADGSEQVITFTQIT--N 251  
QY 101 ASSPIKYNTVETVAVFOYGYKNIRNA-----IEYVGGALNTDTLIF 145  
DB 252 SKKALMTSKDEVSDEKFEFKHIAHMDPLEVIAMKAGTEFYQ-----ALIF 302  
QY 146 TDGELCDVFPVFNADQCELMVKSHY-----KAVPDYCTFVFNVCAXD 190  
DB 303 IPSHAPFLFNSDAKIGMQLYKRVFTMSDCDQMLPMTLRFVKGVDAED 352

RESULT 6  
H71645  
heat shock protein htpg (htpg) RP840 - Rickettsia prowazekii  
C/Species: Rickettsia prowazekii  
C/Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 03-Nov-2000  
C/Accession: H71645  
R.Anderson, S.G.E.; Zomrodipour, A.; Anderson, J.O.; Sichteritz-Ponten, T.; Alsmark, U.  
Nature 396, 133-140, 1998  
A/Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.  
A/Reference number: A71630; MUID:99039499; PMID:9823893  
A/Accession: H71645  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-621 <AND>  
A/Cross-references: GB:AJ235273; GB:AJ235269; NID:g3861237; PIDN:CA15264.1; PID:g386136  
A/Experimental source: strain Madrid E  
A/Genetics:  
A/Gene: htpg; RP840  
A/Superfamily: heat shock protein 90

Query Match 7.5%; Score 84.5; DB 2; Length 621;  
Best Local Similarity 18.9%; Pred. No. 15;  
Matches 39; Conservative 42; Mismatches 88; Indels 37; Gaps 8;  
QY 11 FVSALATQAEITTSKAGENPLMAHE-ELGKY--QDAWKSIDGCVTVYLA--TTYE 65  
DB 127 YSSFWANKEVLTTRKAGSKVHTWESDGLGEIYVADSEQFTGTEIVYIKKSETFL 186  
QY 66 NDTGWSGQFKLCQVQETERRKEDTVTSVTFENASPIKYVYV-----TEYKAVFOY 120  
DB 187 DH-----FLKHIVKSYSDIANPIIFCDEAGNNEQLNSASALMTRPSETED 236  
QY 121 GYKNIRNAIEYQV-----GGLNTDTLIFTDGLCDVFPVFNADQCELMVK 169  
DB 237 QYKEFYKSLSYAVDDPWTLNHNKGALEFTNLIFPESKTFDLFH--PDRKKRVKLYIKR 295  
QY 170 -----SHYKGVDPDYCFVFNVCAXD 190  
DB 296 VFISSDENTILIPSYIRFLRGVDSDD 321

RESULT 7  
AF1394  
autolysin, amidase [imported] - Listeria monocytogenes (strain EGD-e)

C/Species: Listeria monocytogenes  
C/Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C/Accession: AF1394  
R.Glaeser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baguer, F.; Berche, P.; Blo  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duesetzel, O.; Entian, K.D.; Feibi  
D.; Jones, L.M.; Karet, U.  
Science 294, 849-852, 2001  
A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kuzepkat, G.; Madueno, E.; Maitournam, A  
O.K.; C.; Schueter, T.; Simoes, N.; Tixeret, A.; Vazquez-Boland, J.A.; Voss, H.; Weh  
A/Title: Comparative genomes of Listeria species.  
A/Reference number: AB1077; MUID:21537279; PMID:11679669  
A/Accession: AF1394  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-917 <GLA>  
A/Cross-references: GB:NC 003210; PIDN:CAD00636.1; PID:g16412046; GSPDB:GN00177  
A/Experimental source: strain EGD-e  
A/Genetics:  
A/Gene: ami

Query Match 7.5%; Score 84.5; DB 2; Length 917;  
Best Local Similarity 22.6%; Pred. No. 24;  
Matches 37; Conservative 25; Mismatches 49; Indels 53; Gaps 9;

QY 28 GENPLMAHELLGKYQDAWKSIDGCVSVTYVLAATYENDT---GSGV-----S 73  
DB 223 GNGTWSHD-----AVIRYLGITHTDVAIFNQGVFNFNFLINE 265  
QY 74 QFKCLOVEIERKEDTVTSVTFENASPIKY--NTEYKAV--FOYGYKNIR-- 126  
DB 266 KYKAMQV-NVAKIYDRAITAVSRVATGNSVTKENKTEGAKLVNPLSSYSGKNIRII 324  
QY 127 -----NATRYQ--VGGALNTDTLIFTDGLCDVFPVFNADQ 161  
DB 325 REANTSGGTWYQPSIG-----KTIGWVDSKALNTFTYPSMEK 363

RESULT 8  
T12796  
probable transglycoylase - Bacillus subtilis phage SPBc2  
C/Species: Bacillus subtilis phage SPBc2  
C/Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 15-Oct-1999  
C/Accession: T12796; A69911  
R.Izazadeh, V.; Duesetzel, A.; Soldo, B.; Hilbert, H.; Maue, C.; Karamata, D.  
submitted to the EMBL Data Library, August 1997  
A/Description: The complete nucleotide sequence of the Bacillus subtilis SPBc2 pr  
A/Reference number: Z17583  
A/Accession: T12796  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-2285 <LZ>  
A/Cross-references: EMBL:AF020713; NID:g3025478; PID:g3025510; PIDN:AA013005.1  
R.Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, N.M.;  
C.; Bron, S.; Brouillet, S.; Brunsch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,  
Nature 390, 249-256, 1997  
A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Ge  
lech, J.; Harwood, C.R.; Heuvel, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hultio,  
Kostler, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardi  
A/Authors: Lander, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Me  
Y., M.; Ogawa, K.; Ogilvie, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scar  
A/Authors: Schleich, S.; Schroeter, R.; Scollone, F.; Sekiguchi, J.; Sekowski, A.; S  
akuchii, M.; Tamakoshi, A.; Tanaka, T.; Terstegen, P.; Tognoni, A.; Tosato, V.; Uchi  
T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yosh  
A/Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subti  
A/Reference number: A69580; MUID:98044033; PMID:9384377  
A/Accession: A69911  
A/Status: nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-2285 <KUN>  
A/Cross-references: GB:Z59115; GB:AL009126; NID:g2634478; PIDN:CA14053.1; PID:e118;

A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yom1

Query Match 7.4%; Score 83; DB 2; Length 2285;  
Best Local Similarity 23.3%; Pred. No. 1.1e+02;  
Matches 51; Conservative 27; Mismatches 79; Indels 62; Gaps 10;

QY 6 VLLITFSALATQAEITSAKAGENPMAHEILGKYGQAMKSID--GVSTYTYLAKTT 63  
DB 743 LTVSTLVGAFALG-----MALSTLISFAKAKADFEQSQGTNEALTT 790  
QY 64 YENDTSGWSQFKCLQVQIEERKEDYTVTSVTFERNASSPIKYNVTEVAVPQ---Y 120  
DB 791 NKSTDKLIQVYK-----ELQYKESRSLTS-----DEGEYLQVTLQATPALVK 838  
QY 121 GY-----KNINAEIYQVGGGLNTDTLITFDGELCDVFYVPNADQGCGLMVK 169  
DB 839 GYDSQGNAILKTNKELEKALE-----NTKEYLAKKQETPDSAKTPEDSKE--IKK 889  
QY 170 S-----HYKHVPDYCTFVFNVPFCAKDKRTYDIFNEECV 203  
DB 890 SKDELKQYKQIADYND-----KGRPKMDLIADDDY 920

## RESULT 9

C69192  
stomatol-like protein - Methanobacterium thermoautotrophicum (strain Delta H)

C:Species: Methanobacterium thermoautotrophicum  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 24-Sep-1999  
C:Accession: C69192  
R:Smith, D.R.; Doucette-Stamm, L.A.; DeLoughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
Qiu, D.; Spadafora, R.; Viscare, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.  
K.; S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func  
A:Reference number: A69000; MUID:98037514; PMID:9371463  
A:Accession: C69192  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-318 <MTH>  
A:Cross-references: GB:AE000848; GB:AE000666; NID:G2621761; PIDN:AB85197.1; PID:G262177  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH692  
C:Superfamily: erythrocyte band 7 integral membrane protein

Query Match 7.3%; Score 82.5; DB 2; Length 318;  
Best Local Similarity 23.3%; Pred. No. 9.2;  
Matches 38; Conservative 22; Mismatches 77; Indels 25; Gaps 6;

QY 5 VLLITFSALATQAEITSAKAGENPMAHEILGKYGQAMKSIDQSVTYVLAQTT 64  
DB 10 VLVIAFSKILRPYKGVV-----ERLKYQ---RTVSSGLVLIPILEAIK 55  
QY 65 ENDTSGWSQFKCLQVQIEERKEDYTVTSVTFERNASSPIK--YVNTETVAVPOYGYK 123  
DB 56 KVD---WREQYVDPPEVITKNTVVVDCVFYFVVPBNVAVVDFYQALTKMQT 112  
QY 124 NINAEIYQVGGGLNTDTLITFDGELCDVFYVPNADQGCGLM 166  
DB 113 NLRNII-----GDLELDQTL--TSEMINIQLREVDEADTKW 148

## RESULT 10

C97308  
probable acetyltransferase (imported) - Clostridium acetobutylicum  
C:Species: Clostridium acetobutylicum  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
C:Accession: C97308  
R:Noelling, J.; Breton, G.; Omejchko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium

A:Reference number: A96900; MUID:21359325; PMID:21359325  
A:Accession: C97308  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-163 <KUR>  
A:Cross-references: GB:AE001437; PIDN:AA81254.1; PID:G15026402; GSPDB:GN00168  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A:Gene: CAC3322

Query Match 7.3%; Score 92; DB 2; Length 163;  
Best Local Similarity 24.5%; Pred. No. 4.2;  
Matches 37; Conservative 11; Mismatches 51; Indels 52; Gaps 6;

QY 82 EIERKEDYTVTSVTFERNASSPIKYNVTEVAVPQ---FOYGYKNI 125  
DB 3 KISVQMDYETAKQITKNVYKPSIYMMNSDECIRKLLSGYFSAVDKNYIVGYGF 62  
QY 126 RNAIEYQVGGGLNTDTLITFDGELCDVFYVPNADQGCGLMVK 172  
DB 63 GEAQIPVGEQGIYSYKNTDIGILNPNC-----GGIGFDFCCGL----- 107  
QY 173 KAVPYCTFVFNVPFCAKD-RKTYDIFNEECV 202  
DB 108 -----DFRNTLCAKQFRLTVATFNKPAI 131

## RESULT 11

T11876  
hypothetical protein A494R - Chlorella virus PBCV-1

C:Species: Chlorella virus PBCV-1  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T11876  
R:Graves, M.V.; Van Etten, J.L.  
submitted to the EMBL Data Library, May 1999  
A:Reference number: Z18806  
A:Accession: T11876  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-360 <GRA>  
A:Cross-references: EMBL:U42580; NID:G4028896; PIDN:AAC96861.1  
A:Experimental source: specific host Chlorella strain NC64A  
C:Genetics:  
A:Note: A494R

Query Match 7.3%; Score 82; DB 2; Length 360;  
Best Local Similarity 25.8%; Pred. No. 12;  
Matches 41; Conservative 19; Mismatches 59; Indels 40; Gaps 8;

QY 35 HEELGK---YQDAMKSIDQSVTYVLAQTTYENDTSGWSQFKCLQVQIEERKEDYT 91  
DB 43 HEAVNGNIDLPQKQKLENGVN-----TKNIE-----RIKQIEDEINYL 84  
QY 92 VTSVTFERNASSPIKYNVTEV-----KAVPOYGYKNR-NAIE---YQVGGGLNIT 140  
DB 85 LDTI-----PFIEKYVQKETSVDYEQNSVFOYKSNHTNTFRRLTFHVKVSNPT 136  
QY 141 DTLLITFDGELCDVFYVPNADQGCGLMVKKSHYKVPDYC 179  
DB 137 TLDAVDTRETVDQIYCTCGGGMELMKNVNSTGSDLYCNRC 175

## RESULT 12

T11876  
hypothetical protein C03C10.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Sep-2003  
C:Accession: T11876  
R:Berkas, M.  
submitted to the EMBL Data Library, August 1994  
A:Reference number: Z19036  
A:Accession: T11876

A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-381 <WIL>  
A:Cross-references: EMBL:Z35637; PIDN:CAA4688.1; GSPDB:GN00021; CESP:C03C10.3  
A:Experimental source: clone C03C10  
C:Genetics:  
A:Gene: CESP:C03C10.3  
A:Map position: 3  
A:Insertions: 18/3; 137/3; 256/2  
C:Superfamily: ribonucleoside reductase small subunit

Query Match 7.3%; Score 82; DB 2; Length 381;  
Best Local Similarity 21.8%; Pred. No. 13;  
Matches 29; Conservative 20; Mismatches 56; Indels 28; Gaps 3;

QY 9 LTFVSALATQAEITTSKAGENPLMAHEELG-----KYODAMKSIDQGVSVTY 57  
DB 35 LEIVDQTKAASAEETNNESEVNEIDADEPMLQDLDNRFVLPKHHIDINFPYKKAASFM 94  
QY 58 VLAKTYENDTGSWSQKCLQVOEIERKEDYTVSVTFERNASSPI-----KYNV 110  
DB 95 TVEEVDLKDNDM-----EKXNGDEQYFISRLAFPAASDQIVNENICERFSNE 144  
QY 111 TETVKAFFQYGYK 123  
DB 145 VQVSEARFFYGFQ 157

RESULT 13  
D82922  
conserved hypothetical U198 [Imported] - Ureaplasma urealyticum  
C:Species: Ureaplasma urealyticum  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: D82922  
R:Glas, J.I.; Belkowitz, E.J.; Glas, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
submitted to GenBank, February 2000  
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min  
A:Reference number: A82870  
A:Accession: D82922  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-390 <GLA>  
A:Cross-references: GB:AE002119; GB:AF222894; MID:96899153; PIDN:AAF30605.1; GSPDB:GN001  
A:Experimental source: serovar 3; biovar 1  
C:Genetics:  
A:Gene: U198  
A:Genetic code: SGC3

Query Match 7.3%; Score 82; DB 2; Length 390;  
Best Local Similarity 22.3%; Pred. No. 13;  
Matches 29; Conservative 28; Mismatches 51; Indels 22; Gaps 5;

QY 8 LTFVSALATQAEITTSKAGENPLMAHEELGKYODAMKSIDQGVSVTY-VLAKTYEN 66  
DB 167 ILAYKRTGAGANASGAIAEFLTSNPLISKERKKYHIVYIEAG--TYGLIRSTES 223  
QY 67 DTGWSQFKCLQVOEIERKEDYTVSVTFERNASSPIKYVNTETVKAFFQYGY--K 123  
DB 224 NTOAW-----TNISKDNVDSITYL-----STGFLIKVYDQIAKXKIALQNK 268  
QY 124 NIRNAIEYOV 133  
DB 269 DKNVLPPEI 278

RESULT 14  
T08530  
trag protein - Enterobacter aerogenes plasmid R751  
C:Species: Enterobacter aerogenes  
C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: T08530  
R:Thorsted, P.B.; Shah, D.S.; Macartney, D.; Kostelidou, K.; Thomas, C.M.  
plasmid 36, 95-111, 1996

A>Title: Conservation of the genetic switch between replication and transfer genes  
A:Reference number: Z16434; MID:97118926; PMID:8954681  
A:Accession: T08530  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-637 <TMO>  
A:Cross-references: EMBL:U67194; MID:91572520; PIDN:AAC64474.1; PID:91572579  
C:Genetics:  
A:Gene: trag  
A:Genome: plasmid R751

Query Match 7.3%; Score 82; DB 2; Length 637;  
Best Local Similarity 18.3%; Pred. No. 25;  
Matches 46; Conservative 30; Mismatches 89; Indels 86; Gaps 7;

QY 5 VLLITFVSALATQAEITTSKAGE-----NPLMAHEE----- 37  
DB 90 LVSTVGLLGAIVAKVTVTSNNSKANLEYLHSGARBAEKDIOAAGLPRERNVLEIVTKAA 149  
QY 38 --LLGKXQDAMKSID-----QGVSVTVYVLAKTYYENDTG-----SMGQFKCLQY 80  
DB 150 PTATGVYVGGWQDQGNFFYLRHSGPEHVLTYAFTPSGKGVGLVPTLLSMGASSVITDL 209  
QY 81 QEIERKEDYTVSVTFERNASSPIKYVNTETVKAFFQYGYKNIRNAIEYQVGGGLNT 140  
DB 210 -----KGLMALNTAGMFGKNAKKVLRFBPASTSGGVCMNPDEIRLGTGEYEGDVQNTLA 264  
QY 141 DTLIFDTGELCDVFTV-----PNDQGE 164  
DB 265 TLIVDPDGKGLDISHMOKTAPALLVGVILHALYKXKDDGATLPSVDMADLPNRDIG-E 323  
QY 165 LMKKSHYKGV 175  
DB 324 LMMEMATYGHV 334

RESULT 15  
S22992  
trag protein - Escherichia coli plasmid R751

C:Species: Escherichia coli  
C>Date: 31-Dec-1993 #sequence\_revision 02-Aug-1994 #text\_change 08-Oct-1999  
C:Accession: S22992; S42274  
R:Ziegelin, G.; Pansesgrau, W.; Strack, B.; Balzer, D.; Kroeger, M.; Kruff, V.; Lanka  
DNA Seq. 1, 303-327, 1991  
A>Title: Nucleotide sequence and organization of genes flanking the transfer origin  
A:Reference number: S22992; MID:92190548; PMID:165997  
A:Accession: S22992

A:Molecule type: DNA  
A:Residues: 1-637 <ZIE>  
A:Cross-references: EMBL:X54458  
R:Lanka, E.  
submitted to the EMBL Data Library, August 1990  
A:Reference number: S42274  
A:Accession: S42274

A:Molecule type: DNA  
A:Residues: 1-628, 'D', 630-637 <LAN>  
A:Cross-references: EMBL:X54458; MID:942656; PIDN:CAA38927.1; PID:942657  
C:Genetics:  
A:Gene: trag  
A:Genome: plasmid

Query Match 7.3%; Score 82; DB 2; Length 637;  
Best Local Similarity 18.3%; Pred. No. 25;  
Matches 46; Conservative 30; Mismatches 89; Indels 86; Gaps 7;

QY 5 VLLITFVSALATQAEITTSKAGE-----NPLMAHEE----- 37  
DB 90 LVSTVGLLGAIVAKVTVTSNNSKANLEYLHSGARBAEKDIOAAGLPRERNVLEIVTKAA 149  
QY 38 --LLGKXQDAMKSID-----QGVSVTVYVLAKTYYENDTG-----SMGQFKCLQY 80  
DB 150 PTATGVYVGGWQDQGNFFYLRHSGPEHVLTYAFTPSGKGVGLVPTLLSMGASSVITDL 209

Tue Mar 9 14:18:20 2004

us-09-555-296b-4.rpt

```
QY 81 QEI ERKEDYTVTSVFTRNASPIKYNVTETVKA VFOYGYKIRNAIEYQGGJNIT 140
Db 210 -----KGELMALTAGMRQKAKKVLRFEPASTSGGVCWNPIDELRIGTEREVDQNTLA 264
QY 141 DTLIFTDGELCDVFFV-----PNADQCE 164
Db 265 TLIVDPDGKGLDSHQKTA FALIVGVIHLALYKAKDDGGTATLPSVDAMLADPNRDTG-E 323
QY 165 LWTYKSHYKHV 175
Db 324 LMMEMATYGHV 334
```

Search completed: March 9, 2004, 13:34:01  
Job time : 22 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: March 9, 2004, 13:30:56 ; Search time 17 Seconds  
(without alignments)  
640.157 Million cell updates/sec

Title: US-09-555-296B-4  
Perfect score: 1125

Sequence: 1 MKMQVLLLPVSAALATQA.....DRKTYDIFNEECVNGEPWL 209

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 141681 segs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seg length: 0  
Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	290.5	25.8	190	1	HBP2_RHIAP
2	265	23.6	200	1	HBP2_RHIAP
3	252	22.4	190	1	HBP1_RHIAP
4	97.5	8.7	174	1	NUGC_MESVI
5	96.5	8.6	621	1	HTPG_RICCN
6	96	8.5	616	1	HTPG_BORBU
7	87	7.7	639	1	HTPG_TREPA
8	86	7.6	2778	1	FAF_DROME
9	85	7.6	656	1	HTPG_MYCLB
10	84.5	7.5	621	1	HTPG_RICPR
11	82.5	7.3	318	1	V692_MERTH
12	82	7.3	381	1	RIR2_CABEL
13	82	7.3	637	1	TRGS_ECOLI
14	81.5	7.2	627	1	DNAX_RICPR
15	81	7.2	700	1	HS9C_DICDI
16	80.5	7.2	621	1	HTPG_HELPJ
17	79.5	7.1	621	1	HTPG_HELPY
18	79.5	7.1	629	1	HTPG_CHLTP
19	79.5	7.1	725	1	FTKL_ANASP
20	78.5	7.0	864	1	YG3M_YEAST
21	78	6.9	215	1	NO22_SOYBN
22	77	6.8	523	1	CP5C_CANMA
23	77	6.8	909	1	HEX_ADEMI
24	76.5	6.8	393	1	IDH_STRMU
25	76.5	6.8	724	1	DDX4_HUMAN
26	76	6.8	260	1	DKC_MOUSE
27	76	6.8	1031	1	YDG9_SCHPO
28	76	6.8	1066	1	HYSA_STRPN
29	76	6.8	1068	1	TRI_SULSO
30	75.5	6.7	1489	1	RNG2_SCHPO
31	75	6.7	414	1	Y878_MERJA
32	75	6.7	424	1	PORA_CAMJE
33	75	6.7	566	1	BCDO_RAT

34	75	6.7	626	1	HTPG_BURBP
35	75	6.7	647	1	HTPG_MTCU
36	75	6.7	6629	1	R1AB_IBVC
37	75	6.7	6629	1	R1AB_IBVC
38	74.5	6.6	400	1	GTR3_RABIT
39	74.5	6.6	492	1	YGCE_ECOLI
40	74.5	6.6	726	1	NP11_YEAST
41	74.5	6.6	890	1	LPN1_HUMAN
42	74.5	6.6	2768	1	THYG_HUMAN
43	74	6.6	406	1	IDH_SPTA
44	74	6.6	782	1	POT6_ARATH
45	74	6.6	854	1	TRIC_SULTO

ALIGNMENTS

RESULT 1					
HBP2_RHIAP	STANDARD;	PRT;	190 AA.		
AC	O77421;				
DT	30-MAY-2000 (Rel. 39, Created)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Female-specific histamine-binding protein 2 precursor (FS-HBP2).				
OS	Rhipicephalus appendiculatus (Brown ear tick).				
OC	Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;				
OC	Parasitiformes; Ixodidae; Ixodidae; Rhipicephalus.				
OX	NCBI_TaxID=34631;				
RP	SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (1.25 ANGSTROMS).				
RX	TISSUE=Salivary gland;				
RC	MEDLINE=99289454; PubMed=10360182;				
RA	Paesen G.C., Adams P.L., Harlos K., Nuttall P.A., Stuart D.I.;				
RT	"Tick histamine-binding proteins: isolation, cloning, and three-dimensional structure".				
RL	Mol. Cell 3:661-671 (1999).				
CC	-1- FUNCTION: Binds histamine with a high-affinity. The ability to				
CC	outcompete histamine receptors indicates that its function is to				
CC	suppress inflammation during blood feeding.				
CC	-1- SUBCELLULAR LOCATION: Secreted.				
CC	-----				
CC	EMBL; U96081; AAC63107.1; -.				
DR	PDB; 1QFT; 19-APR-00.				
DR	PDB; 1QFW; 19-APR-00.				
DR	InterPro; IPR002970; His_binding.				
DR	Pfam; PF02098; His_binding; 1.				
DR	ProDom; PD15455; His_binding; 1.				
KW	Signal; 3d-structure.				
FT	SIGNAL	1	19		
FT	CHAIN	20	190		
FT					
FT	DISULFID	67	188		
FT	DISULFID	138	167		
FT	TURN	23	24		
FT	TURN	27	30		
FT	HELIIX	31	33		
FT	HELIIX	36	41		
FT	TURN	42	45		
FT	TURN	48	53		
FT	STRAND	57	58		
FT	TURN	59	61		
FT	STRAND	62	63		
FT	TURN	64	64		
FT	STRAND	66	76		

-----

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

FEMALE-SPECIFIC HISTAMINE-BINDING PROTEIN



FT CHAIN 19 190 FEMALE-SPECIFIC HISTAMINE-BINDING PROTEIN  
 FT DISULFID 66 187 1. SIMILARITY.  
 FT DISULFID 137 166 BY SIMILARITY.  
 SQ SEQUENCE 190 AA, 21370 MW, 855BE151A90053B1 CRC64;

Query Match 22.4%; Score 252; DB 1; Length 190;  
 Best Local Similarity 33.5%; Pred. No. 2.6e-15;  
 Matches 68; Conservative 28; Mismatches 89; Indels 18; Gaps 8;

QY 3 NOVVLITFVSALATQAEITTSKAGENPLMAHEELGKYDANKSIDQGVSVTYLAKT 62  
 DB 1 MKLLSLIAFVLALSGVKA-----DKPWADEKANGEHQDAWKHLQKVEENYDLIKA 52  
 QY 63 TYENDGWSGQFKLCQVEIERKEEDTYTTSVTFERRASSPICYNVTEVYKAVPQYGY 122  
 DB 53 IYKNDP-VKGNDFTCVGRDAQNLNEDENKXWEMFNKNAUTV-YQHTFEKATPDKNYGY 110  
 QY 123 KQIRNAIEYQVGGINTDTLTFTDGLCDVFYYPNAD---QGCETLVKSKHYKRPDYC 179  
 DB 111 -KKNATYQTEDEGVLHDVLAFSQ-DNCVITYALGPDGSGAGYELMA--TDYTDVPASC 166  
 QY 180 TEVENVECAKOKRTDINEBCV 202  
 DB 167 LEKNEYDA-GLPVRDVTISDCL 188

RESULT 4  
 NUOG MESVI STANDARD; PRT; 174 AA.  
 AC Q9MRL:  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE NAD(P)H-quinone oxidoreductase chain J, chloroplast (EC 1.6.5.-)  
 DE (NAD(P)H dehydrogenase, chain J) (NADH-plastoquinone oxidoreductase subunit J).  
 GN NDHJ.  
 OS Mesostigma viride.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Mesostigmatophyceae;  
 CC Mesostigmatales; Mesostigmataceae; Mesostigma.  
 CX NCBI\_TaxID=41882;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NIES-296;  
 RX MEDLINE=20150907; PubMed=10688199;  
 RA Lemieux C., Orlis C., Turmel M.;  
 RT "Ancestral chloroplast genome in Mesostigma viride reveals an early branch of green plant evolution.";  
 RT Nature 403:649-652(2000).  
 RL Nature 403:649-652(2000).

CC -1- CATALYTIC ACTIVITY: NAD(P)H + plastoquinone = NAD(P) (+) + plastoquinol.  
 CC -1- SIMILARITY: Belongs to the complex I 30 kDa subunit family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF166114; AAF43839.1;  
 DR InterPro: IPR001268; Complex1\_30K.  
 DR Pfam: PF00329; complex1\_30kd; 1.  
 DR PROSITE: PD001581; Complex1\_30K; 1.  
 DR PROSITE: PS00542; COMPLEX1\_30K; 1.  
 KW Oxidoreductase; NAD; NADP; Quinone; Plastoquinone; Chloroplast.  
 SQ SEQUENCE 174 AA, 20329 MW, 122ABFA918BD47B7 CRC64;

Query Match 8.7%; Score 97.5; DB 1; Length 174;  
 Best Local Similarity 27.0%; Pred. No. 0.11;

Matches 31; Conservative 20; Mismatches 37; Indels 27; Gaps 5;

QY 105 IKYNNVETKAVPQYGYKIRNAIEYQVGGINTDTLTFTDGLCDVFYV---BNAD 160  
 DB 47 VQACMLTSVALALYAGFNLYRSQCAVDVSPG-----GDLASVHLTFVDDNAD 95

QY 161 QGCETLVKSKHYKRPDYCTFVFNVPCA-----KQKRTDINEBCVYGEPL 209  
 DB 96 QPEVCIR---VVPRIKPIIPEVFWIMKTADQENESIDMG--IYEGHPL 144

RESULT 5  
 HTPG\_RICCN STANDARD; PRT; 621 AA.  
 ID HTPG\_RICCN  
 AC P58478;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Chaperone protein htpg (Heat shock protein htpg) (High temperature protein G).  
 DE HTPG OR RC1302.  
 GN Rickettsia conorii.  
 OS Rickettsia conorii.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;  
 CC Rickettsiaceae; Rickettsiae; Rickettsia.  
 CX NCBI\_TaxID=781;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Malish 7;  
 RX MEDLINE=2142074; PubMed=11557893;  
 RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V., Samson D., Roux V., Coesset P., Weissenbach J., Claverie J.-M., Raoult D.;  
 RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";  
 RT Science 293:2093-2098(2001).  
 RL -1- FUNCTION: Molecular chaperone. Has ATPase activity (By similarity).  
 CC -1- SUBUNIT: Homodimer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: Belongs to the heat shock protein 90 family.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AE008676; AL003840.1;  
 DR PIR: P97862; F97862.1;  
 DR HAMAP: MF\_00505; 1.  
 DR InterPro: IPR003594; ATPbind\_ATPase.  
 DR InterPro: IPR001404; Hsp90.  
 DR Pfam: PF02518; HATPase\_C; 1.  
 DR Pfam: PF00183; HSP90; 2.  
 DR PRINTS: PR00775; HEATSHOCK90.  
 DR SMART: SM00387; HATPase\_C; 1.  
 DR PROSITE: PS00298; HSP90; 1.  
 KW Chaperone; ATP-binding; Heat shock; Complete proteome.  
 FT DOMAIN 1 328  
 FT DOMAIN 329 544 A; SUBSTRATE-BINDING (BY SIMILARITY).  
 FT DOMAIN 545 621 B (BY SIMILARITY).  
 SQ SEQUENCE 621 AA, 70804 MW, FFE5B352DD8F017 CRC64;

Query Match 8.6%; Score 96.5; DB 1; Length 621;  
 Best Local Similarity 21.0%; Pred. No. 0.63;

Matches 43; Conservative 46; Mismatches 81; Indels 35; Gaps 10;

QY 11 FVSALATQAEITTSKAGENP-LMAHEELGKY-QDAWKSIDQGVSVTYLAKTYEN 66  
 DB 127 YSSFMVADKVTYTSRKXGDEKVIW-ESDGLGYTSDSDKEFTGTETVLHK---BE 182  
 QY 67 DTGWSGQFKLCQVEIERKEEDTYTTSVTF-----RNASSPICYNVTEVYKAV 117

Db 183 DT--FLDHF--LKHIVSYSDHIAVPIYFDEAGNNEIQLNSASALMTRPKSEITEEQ 237  
 QY 118 FOYGNRNRAIE-----YQYGGGLNTDILFIDGELCDVFVFNADGCELMVK- 169  
 Db 238 YKEFYKSLSYALDDPMTIMANKKEGALFEFTNLFISSKTFPLFH-PDKRKYVLYIKRV 296  
 QY 170 ----SHYKVPDYCTFVNFVCAKD 190  
 Db 297 FISDENIDILPSYLRFLRGVDSED 321

## RESULT 6

HTPG BORBU STANDARD; PRT; 616 AA.

AC P42555;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Chaperone protein htpg (Heat shock protein htpg) (High temperature protein G).  
 DE HTPG OR B0560.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.  
 CX NCBI\_Taxid=139;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RA Strain=297;  
 RA Porcella S.F., Radolf J.D., Norgard M.V.;  
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBCC databases.  
 RN [2]  
 RC SEQUENCE FROM N.A.  
 RA STRAIN=ATCC 35210 / B31;  
 RC MEDLINE=98065943; PubMed=9403685;  
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,  
 RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,  
 RA Dougherty B., Tomb J.F., Fleischmann R.D., Richardson D.,  
 RA Peterson J., Kervatage A.R., Quackenbush J., Salzberg S., Hanson M.,  
 RA Van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman C.,  
 RA Ureback T., Matthey L., McDonald L., Artach P., Bowman C.,  
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,  
 RA Smith H.O., Venter J.C.;  
 RT "Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi."  
 RL Nature 390:580-586(1997).  
 RN [3]  
 RC SEQUENCE OF 497-574 FROM N.A.  
 RP STRAIN=212;  
 RC MEDLINE=9511614; PubMed=7812434;  
 RA Ojaimi C., Davidson B.E., Saint-Girons I., Old I.G.;  
 RT "Conservation of gene arrangement and an unusual organization of rRNA genes in the linear chromosomes of the Lyme disease spirochaetes  
 RT Borrelia burgdorferi, B. garinii and B. atzei.";  
 RL Microbiology 140:2931-2940(1994).  
 RN [1]  
 RC FUNCTION: Molecular chaperone. Has ATPase activity (By similarity).  
 CC -1- SUBUNIT: Homodimer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: Belongs to the heat shock protein 90 family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: U51878; AAA97469.1;  
 DR EMBL: AE001157; AAC66919.1; ALT\_INT.  
 DR EMBL: U32145; AAC41403.1;  
 DR HSSP: P07900; IYER.  
 DR TIGR: BB0560; --

DR HAMAP; MF\_00505; --; 1.  
 DR InterPro; IPR003594; ATPbind\_ATPase.  
 DR InterPro; IPR001404; Hsp90.  
 DR Pfam; PF02518; HATPase\_C; 1.  
 DR Pfam; PF00183; HSP90\_2.  
 DR PRINTS; PR00775; HEATSHOCK90.  
 DR SMART; SM00387; HATPase\_C; 1.  
 DR PROSITE; PS00298; HSP90\_1.  
 KM Chaperone; ATP-binding; Heat shock; Complete proteome.  
 FT DOMAIN 1 333 A; SUBSTRATE-BINDING (By similarity).  
 FT DOMAIN 334 542 B (By similarity).  
 FT DOMAIN 543 616 C.  
 SQ SEQUENCE 616 AA; 71218 MW; D676552F48DBE84 CRC64;

## Query Match

Best Local Similarity 21.5%; Score 96; DB 1; Length 616;  
 Matches 46; Conservative 34; Mismatches 92; Indels 42; Gaps 8;

QY 11 FVSALATQ-AETSAKAGNP--LMAHELLGKYQDAWKSIDGVSVTVYLAATTEND 67  
 Db 121 FVSALFSEKVEVTSKALLSDPAYIWSDGKGYEIEKAKESGTEIKLYINKEGLE-- 178  
 QY 68 TGSWSQSFCKLOY-----QIEKREKDYTVTSFTPRNA 101  
 Db 179 --YANKWKIOIIRKYSNHNINYPYIKYSPIMKDGKQBIIEKEKELMTALMTKN- 234  
 QY 102 SSPIKYVNTETVKAVFQYGYGNINRAIEYQVGGGLNTDILFIDGELCDVFVFNADQ 161  
 Db 235 KSEITAEENEFYKNT-TPDYENPLMHTAEGLEVTNLPYPSKAPVLYX-PTMKP 292  
 QY 162 GCELMVK-----SHYKVPDYCTFVNFVCAKD 190  
 Db 293 GVTALFNRIFITDSEGLPNTLRFIKGIIIDCOD 326

## RESULT 7

HTPG TREPA STANDARD; PRT; 639 AA.

AC O83949;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Chaperone protein htpg (Heat shock protein htpg) (High temperature protein G).  
 DE HTPG OR TP0984.  
 GN Treponema pallidum.  
 OS Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.  
 CX NCBI\_Taxid=160;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=Nichols;  
 RC MEDLINE=98332770; PubMed=9665876;  
 RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,  
 RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,  
 RA Sodergren E., Hardam J.M., McLeod M.P., Salzberg S., Peterson J.,  
 RA Kralak H., Richardson D., Howell J.K., Chidambaram M., Ureback T.,  
 RA McDonald L., Artach P., Bowman C., Cotton M.D., Weidman J., Garland S.,  
 RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,  
 RA Venter J.C.;  
 RT "Complete genome sequence of Treponema pallidum, the syphilis spirochete."  
 RL Science 281:375-388(1998).  
 RN [1]  
 RC FUNCTION: Molecular chaperone. Has ATPase activity (By similarity).  
 CC -1- SUBUNIT: Homodimer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: Belongs to the heat shock protein 90 family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial



CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: AE001265; AAC65938.1; -

CC PIR: F71258; F71258.

DR HSSP: F02829; 1A88.

DR TIGR: TP0984; -

DR HAMAP: MF\_00505; -; 1.

DR InterPro: IPR003594; Atgbind\_ATPase.

DR InterPro: IPR001404; Hsp90.

DR Pfam: PF02518; HATPase\_C; 1.

DR Pfam: PF00183; HSP90; 2.

DR PRINTS: PR00775; HEATSHOCK90.

DR SMART: SM00387; HATPase\_C; 1.

DR PROSITE: PS00298; HSP90; 1.

KM Chaperone; ATP-binding; Heat shock; Complete proteome.

FT DOMAIN 1 348 A; SUBSTRATE-BINDING (BY SIMILARITY).

FT DOMAIN 349 565 B (BY SIMILARITY).

FT DOMAIN 566 639 C (BY SIMILARITY).

SQ SEQUENCE 639 AA; 72937 MW; 388FDBAC2282C31D CRC64;

Query Match 7.7%; Score 87; DB 1; Length 639;

Best Local Similarity 22.9%; Pred. No. 4.5;

Matches 50; Conservative 26; Mismatches 52; Indels 90; Gaps 10;

QY 5 VVLLTFYSALAAQAEITSAKAGENPLMAHEELGKYD-----AMKSIDGVIS 54

DB 182 VVLLHSQENSEFAR-----MLEEYTKISDHPYLYHLQKEYDKDA 228

QY 55 VTYVLAKTYYENDTSGMGSQPKCLQVOEIRKEDY-----TYSVFTRNASSPIKYN 109

DB 229 VTDQKQKVDQVNDGALMKRKS-----ELKEDYRFRYQTLT-----RDSFPLLYVH 277

QY 110 VTEIVKAVFYQYGYKINRNALISYQGGSLNTDITLFDG--ELCDVFYYP-----NA 159

DB 278 TK-----AEGQEVYVTLFYVPAPAKAPFDLFA 303

QY 160 D--OGCELMVKK-----SHYGVVDYCTFVFNVCARD 190

DB 304 DYKPGVKLFVKRVPITTDKELLYVILFRGVIDSBD 341

RESULT 8

PAF\_DROME STANDARD; PRT; 2778 AA.

AC P55824; Q9Y9T6; Q9Y027;

DT 01-NOV-1997 (Rel. 35, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Probable ubiquitin carboxyl-terminal hydrolase PAF (EC 3.1.2.15)

DE (Ubiquitin thiolesterase PAF) (Ubiquitin-specific processing protease PAF) (Deubiquitinating enzyme PAF) (Fat facets protein).

GN PAF OR BCDNA:JD2582 OR CG1945.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;

XP [1]

RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 3), AND TISSUE SPECIFICITY.

RC TISSUE=Eye, imaginal disk; PubMed=1295747;

RX MEDLINE=93202020; PubMed=1295747;

RA Fischer-Vize J.A., Rubin G.W., Lehmann R.,

RT "The fat facets gene is required for Drosophila eye and embryo development.";

RT development 116:985-1000(1992).

RP [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkley;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amandites P.G., Scherer S.E., Li P.W., Hopkins R.A., Galie R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Change M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abil J.F., Abdayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,

RA Ballew R.M., Bastru A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Borchman M.R., Bouck J., Brokstein P., Brottier P.,

RA Butris K.C., Buean D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,

RA de Faliros B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,

RA Folser C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glaser K.,

RA Glodet A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.A.,

RA Jatali M., Kalush F., Kapten G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Laeko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Melnikov G., Mishina N.V., Mobarry C., Morris J.C., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Munz D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S.-Y., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang Q., Zhao Q., Zheng L.,

RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Zhang G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.";

RT Science 287:2185-2195(2000).

RN [3]

RP REVISIONS, AND ALTERNATIVE SPLICING.

RX MEDLINE=22426069; PubMed=12537572;

RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,

RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochuk S.B.,

RA Smith C.D., Tupy J.L., Whitfield J.S., Bayraktaroglu L., Berman B.P.,

RA Bettencourt B.R., Celinker S.E., de Grey A.D.N.J., Drysdale R.A.,

RA Harris N.L., Richter U., Russo S., Schroeder A.J., Shu S.Q.,

RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,

RA Lewis S.E.;

RT "Annotation of the Drosophila melanogaster euchromatic genome: a systematic review.";

RT Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).

RN [4]

RP SEQUENCE OF 1089-2778 FROM N.A. (ISOFORM 1).

RC STRAIN=Berkley;

RX MEDLINE=20196012; PubMed=10731138;

RA Rubin G.M., Hong L., Brokstein P., Evans-Holm M., Frise E.,

RA Stapleton M., Harvey D.A.;

RT "A Drosophila complementary DNA resource.";

RT Science 287:2222-2224(2000).

RP [5]

RP FUNCTION: Required for eye and embryo development, and plays a

role in compound eye assembly and oogenesis respectively. In the

larval eye disks, cells outside the assembling facets require this

protein for short-range cell interactions that prevent the myosin

cells from becoming photoreceptors. It is also required for

nuclear migration and cellularization in early embryogenesis and

could play a role in pole cell determination, development or

function.

RP [6]

RP CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =

ubiquitin + a thiol.

RP [7]

RP ALTERNATIVE PRODUCTS:

Event=Alternative splicing; Named isoforms=3;

Comment=Experimental confirmation may be lacking for some

isoforms;

Name=1;

Isoid=P55824-1; Sequence=Displayed;

Name=2;

	CC	-- isoId=P55824-2; Sequence=VSP_005270;
	CC	Name=3;
	CC	-- isoId=P55824-3; Sequence=VSP_005269;
	CC	-- TISSUE SPECIFICITY: Eye disks and ovaries.
	CC	-- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.
	CC	- - SIMILARITY: Belongs to peptidase family C19.
	CC	-----
	CC	This SMS-S-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
	CC	-----
Dc	DR	EMBL; L04959; AAF01345.1; -
Dc	DR	EMBL; L04958; AAF01346.1; -
Dc	DR	EMBL; L04960; AAF01347.1; -
Dc	DR	EMBL; L04960; AAF01348.1; -
Dc	DR	EMBL; AB003779; AA57198.1; -
Dc	DR	EMBL; AB003779; AA14291.1; -
Dc	DR	EMBL; AF145677; AAC38652.1; -
Dc	DR	MEROPE; C19_007; -
Dc	DR	Flybase; FBgn005632; faf.
Dc	DR	GO; GO:0005737; Cytoplasm; IDA.
Dc	DR	GO; GO:0007349; Piceallularization; IMP.
Dc	DR	GO; GO:0007995; Plembyonic morphogenesis; IMP.
Dc	DR	GO; GO:0007456; Peye morphogenesis (sensu Drosophila); IMP.
Dc	DR	GO; GO:0008583; Pmyctery cell fate differentiation (sensu Dr. .); IMP.
Dc	DR	GO; GO:0007097; Pnuclear migration; IMP.
Dc	DR	GO; GO:0016579; Proteine dubiquitination; IDA.
Dc	DR	GO; GO:0006511; Pubilitin-dependent protein catabolism; IGI.
Dc	DR	InterPro; IPRO01394; Peptide_C19.
Dc	DR	Pfam; PF00443; UCH_1.
Dc	DR	PROSITE; PS00972; UCH_2_1; 1.
Dc	DR	PROSITE; PS00973; UCH_2_2; 1.
Dc	DR	PROSITE; PS02035; UCH_2_3; 1.
Kw	KW	Ubl conjugation pathway; Hydrolase; Thiol protease;
Kw	KW	Developmental protein; Vision; Alternative splicing.
FT	FT	ACT_SITE 1677 1678
FT	FT	BY SIMILARITY.
FT	FT	ACT_SITE 1978 1978
FT	FT	EX SIMILARITY.
FT	FT	BX SIMILARITY.
FT	FT	KCRRTIKKLVSQDEBDATTAATTAAITVTTSPTATA
FT	FT	ALEPAGSGLTTVENKLLISQENPOAKSLD -> VTRA
FT	FT	NNV (in isoform 3).
FT	FT	/Ffid=VSP_005269.
FT	FT	IATATTEBPGMSLTMTVMKNLIIGENPOAKSLD ->
FT	FT	SOROL (in isoform 2).
FT	FT	/Ffid=VSP_005270.
FT	FT	CONFLICT 234 234
FT	FT	E -> D (IN REF. 1) , AAF01345).
FT	FT	CONFICT 2725 2725
FT	FT	T -> S (IN REF. 1)
SQ	SQ	SEQUENCE 2778 AA; 311139 MW; FFB90438BA53A02B CRC64;
Query March		7.6%; Score 86; DB: 1; Length 2778;
Beef Local Similarity		22.9%; Pred. No. 33; Indels 82; Gaps 12
Matches		54; Conservative 28; Mismatch 72;
Oy	6	VILLTVASALATOETTSAKAGENPLMAHEELG-----KYODA-----WKSIDOG 52
Dd	1756	VILTKHQAIPAHLGHSAIQYVVRGLTWTHFLKGEPVNLRBOODAVEFMSTLESLEDG 181
Oy	53	VASTTYVLAKTYEND--GSNSQPKCQOVEIERKEEDYTYSVFIFPRASPPIRYNV 110
Dd	1816	LK--ALGGPOLMNATLGSSFSDDQICQECPHRYSKEPSVSV-DIRNHS-----L 186
Oy	111	TEVVAVFGYGNIINAIEYOVGGIINTDTLLIFTGTGE---CDVFFVPNAOQCELV 167
Dd	1866	TESLDQ-----YVGELLEGAADVAHQDKVKV-VIV 189
Oy	168	KSHHYKAVP-----DY---CFYFNVRPCADRKYDI PNDECYNSEPM 208
Dd	1897	KRCVCVKLPVLAIQLREFEYERVCAIKRN-----DFEFPRLIDNEPY 1942

```

01 137 YSFPVANKVVELTRKAGETATATRSWD--GEATYTIYSVDAPQGTSTYTHLHKEDDF 193
02 66 NDGSGSGQGFCKLQV-----QETLRK-----EDYTVSVTFPRN 100
03
04
05
06
07
08
09
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

```

Db 194 DELHDYTSWKRLVKKYKSDPLAMPIMEVERBARAPATSDGADGSEQYTIQNTI--N 251

QY 101 ASSPKRYNNVETKAVFQYGYKNIRNA-----IEYVGGGLNTDTLIF 145

Db 252 SKMALMTKSKDEVSDEYKEFEKHAHAMDDPLEVIAMKAEGRFEYQ-----ALLF 302

QY 146 TDGELCDVFYVYNADOGCELMWYKSHY-----KAVPDYCTVFVNFCAKD 190

Db 303 IFSHAPFDLFSNDAKIGMQLYKRVFIMSDCDQMLPMYLRKRVKVAAD 352

RESULT 10

HTPG\_RICPR STANDARD; PRT; 621 AA.

AC 09ZCB9;

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Chaperone protein htpg (Heat shock protein htps) (High temperature protein G).

GN HTPG OR Rp840.

OS Rickettsia prowazekii.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;

OC Rickettsiaceae; Rickettsiae; Rickettsia.

OX NCBI\_TaxID=782;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Madrid E;

RX MEDLINE=99039499; PubMed=9823893;

RA Andersson S.G.E., Zomrodipour A., Andersson J.O.,

RA Scharitz-Ponten T., Almaraz U.C.M., Podowski R.W., Naeslund A.K.,

RA Eriksson A.-S., Winkler H.H., Kurland C.G.;

RT "The genome sequence of Rickettsia prowazekii and the origin of mitochondria."

RT Nature 396:133-140 (1998).

RL Nature 396:133-140 (1998).

CC -1- FUNCTION: Molecular chaperone. Has ATPase activity (By similarity).

CC -1- SUBUNIT: Homodimer (By similarity).

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -1- SIMILARITY: Belongs to the heat shock protein 90 family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: AJ235273; CA15264.1; -

DR PIR: H71645; H71645.

DR HSSP: P02829; 1A4H.

DR HAMAP: MF\_00505; -; 1.

DR InterPro: IPR003594; ATPbind\_ATPase.

DR InterPro: IPR001404; Hsp90.

DR Pfam: PF02518; HATPase\_c; 1.

DR Pfam: PF00183; HSP90\_2.

DR PRINTS: PR00775; HEATSHOCK90.

DR SMART: SM00387; HATPase\_c; 1.

DR PROSITE: PS00298; HSP90; 1.

DR Chaperone; ATP-binding; Heat shock; Complete proteome.

FT DOMAIN 1 328 A; SUBSTRATE-BINDING (BY SIMILARITY).

FT DOMAIN 329 544 B (BY SIMILARITY).

FT DOMAIN 545 621 C.

FT SEQUENCE 621 AA; 70713 MW; 9f97b97801524007 CRC64;

QY Query Match 7.5%; Score 84.5; DB 1; Length 621;

Best Local Similarity 18.9%; Pred. No. 7.3;

Matches 39; Conservative 42; Mismatches 88; Indels 37; Gaps 8;

QY 11 PYSAALATQAEFTTSKAGENPLMAHE-ELGKY--QDAKSIDGVSVTVYLA--TTYE 65

Db 127 YSSFMVADKVTVYTSKAGSKVHTWESDGLGEIYVADSEQFTGTETIVYIKKSETTEL 186

QY 66 NDTSWGSQFKCLQVOEIERKEEDYTVTSVFTRNASSPKRYNV-----TEVKAIVFX 120

Db 187 DH-----FLKRIIVASDIAVVFYPCDEAGNNEIQUNASALMTREKSEITED 236

QY 121 GYKNIRNAIEYQVG-----GGANTDTLIFTDGELCDVFYVYNADOGCELMWVK 169

Db 237 QYKERYKSLSAVDDPMVTLNNKGAIEFTNLLFISSTFDLPH-PDRKAVKLYIKR 295

QY 170 -----SAKRVPDYCTVFVNFCAKD 190

Db 296 VFISDENIDLIPTSYLRFLRGVISED 321

RESULT 11

Y692\_METTH STANDARD; PRT; 318 AA.

AC 026788;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein MTH692.

GN MTH692.

OS Methanobacterium thermoautotrophicum.

OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;

OC Methanobacteriaceae; Methanothermobacter.

OX NCBI\_TaxID=187420;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Delta H;

RX MEDLINE=98037514; PubMed=9371463;

RA Smith D.R., Doucette-Stamm L.A., Delonghery C., Lee H.-M., Dubois J.,

RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,

RA Harrison D., Hoang L., Keagle P., Lumm W., Porlier B., Qiu D.,

RA Spadafora R., Viscare R., Wang Y., Wierzbowski J., Gibson R.,

RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,

RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,

RA Daniels C.J., Mao J.-T., Rice P., Noelling J., Reeve J.N.,

RT "Complete genome sequence of Methanobacterium thermoautotrophicum J. Bacteriol. 179:7135-7155 (1997).

RL J. Bacteriol. 179:7135-7155 (1997).

CC -1- SIMILARITY: Belongs to the band 7 / mec-2 family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: AB000848; AAB85197.1; -

DR PIR: C69192; C69192.

DR InterPro: IPR001107; Band\_7.

DR InterPro: IPR001972; Stomatlin.

DR Pfam: PF01145; Band\_7; 1.

DR PRINTS: PR00721; STOMATLIN.

DR SMART: SM00244; PHB; 1.

DR PROSITE: PS01270; BAND\_7; 1.

DR Hypothetical protein; Transmembrane; Complete proteome.

FT TRANSMEM 2 22 POTENTIAL.

FT SEQUENCE 318 AA; 35425 MW; 4A72C0A08E99278D CRC64;

QY Query Match 7.3%; Score 82.5; DB 1; Length 318;

Best Local Similarity 23.3%; Pred. No. 4.9;

Matches 38; Conservative 23; Mismatches 77; Indels 25; Gaps 6;

QY 5 VVLLTFYSAALATQAEFTTSKAGENPLMAHEELGKYQDAKSIDGVSVTVYLA--TTYE 64

Db 10 VLVVLAFLSKILRLPYEYGVV-----ERLGRQ--RTVESGLVVIIPFEAIK 55

QY 65 ENDIGSWSSQKCLQVOEIERKEEDYTVTSVFTRNASSPKRYNV-----TEVKAIVFX 123

```

Db      56 KVD---MEQVVDVPEQVITKNTVVVDVCFFVEYVDFEFAVNVVDFYQALITKAC 112
Qy      124 NRNAIEYOVGGGINTDTLFTDGLCDVFFVFNADQGCGLW 166
Db      113 NLRNII-----GDLELDQTL--TSREWINTQLREVLEADTKW 148

RESULT 12
RIR2_CAEEL STANDARD; PRT; 361 AA.
AC P42170;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ribonucleoside-diphosphate reductase small chain (EC 1.17.4.1)
DE (Ribonucleoside reductase).
GN RNR-2 OR C03C10.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Berk's M.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Provides the precursors necessary for DNA synthesis.
CC -1- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized
CC thiorodoxin + H(2)O = ribonucleoside diphosphate + reduced
CC thiorodoxin.
CC -1- COFACTOR: Binds 2 iron ions per subunit (By similarity).
CC -1- PATHWAY: DNA replication pathway; first step.
CC -1- SUBUNIT: Heterodimer of a large and a small chain.
CC -1- SIMILARITY: Belongs to the ribonucleoside diphosphate reductase
CC small chain family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z35637; CAA84688.1; -.
DR PIR; T18876; T18876.
DR HSRP; P1157; 1XSM.
DR WormPep; C03C10.3; CE00874.
DR InterPro; IPR000358; Ribonuc_redctase.
DR Pfam; PF00268; ribnuc_red_sml_1.
DR PROSITE; PS00368; RIBONUC_SMALL; 1.
KW Oxidoreductase; DNA replication; Metal-binding; Iron.
FT METAL 130 130 IRON 1 (BY SIMILARITY).
FT METAL 161 161 IRON 1 AND 2 (BY SIMILARITY).
FT METAL 164 164 IRON 1 (BY SIMILARITY).
FT METAL 224 224 IRON 2 (BY SIMILARITY).
FT METAL 258 258 IRON 2 (BY SIMILARITY).
FT METAL 261 261 IRON 2 (BY SIMILARITY).
FT ACT_SITE 168 168 BY SIMILARITY.
SQ SEQUENCE 361 AA; 44289 MW; 75497147ABF6C59 CRC64;

Query March 7.3%; Score 82; DB 1; Length 361;
Best Local Similarity 21.8%; Pred. No. 6.7;
Matches 29; Conservative 20; Mismatches 56; Indels 28; Gaps 3;

```

```

Qy      111 TETKAVFQYQYK 123
Db      145 VQVSEARFFYGFQ 157

RESULT 13
TRGS_ECOLI STANDARD; PRT; 637 AA.
AC Q00184;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Conjugal transfer protein trag.
GN TRAG.
OS Escherichia coli.
OC Plasmid Incp-beta R751.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HB101;
RX MEDLINE=92190548; PubMed=1665997;
RA Ziegellin G., Pansegr W., Strack B., Balzer D., Kroege M.,
RA Kruff V., Lanka E.;
RT "Nucleotide sequence and organization of genes flanking the transfer
RL DNA seq. 1:303-327(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Thomas C.M.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Required for conjugative transfer of plasmid R751.
CC Binds tightly and specifically to the relaxase tral. Can also bind
CC to DNA without sequence specificity. May form a pore-like
CC structure that could serve as a channel for DNA transfer (By
CC similarity).
CC -1- SUBUNIT: May form multimers of at least 18 subunits (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (By similarity).
CC -1- SIMILARITY: BELONGS TO THE VIRP4/TRAG FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X54458; CAA38327.1; -.
DR EMBL; U67194; AAC64474.1; -.
DR InterPro; IPR003688; TRAG.
DR Pfam; PF02534; TRAG; 1.
FT PLASMIID; Conjugation; DNA-binding; Transmembrane; Inner membrane.
FT DOMAIN 1 22 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 23 43 POTENTIAL.
FT DOMAIN 44 84 POTENTIAL.
FT TRANSMEM 85 105 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 106 637 POTENTIAL.
SQ SEQUENCE 637 AA; 69883 MW; 7B45A9A9020902FB CRC64;

Query March 7.3%; Score 82; DB 1; Length 637;
Best Local Similarity 18.3%; Pred. No. 12;
Matches 46; Conservative 30; Mismatches 89; Indels 86; Gaps 7;

```

Db 150 PRTATGYVVGWQKDKNFFYLHSHGPEHVLTAPTRSKGGVGLVPTLLISMGASVITDL 209  
 QY 81 QEIERKEEDYIVSVETFPNNSPFIKYVAVETVYAVFYQYGVKINRNAREVQVGGNIT 140  
 Db 210 -----KGEIMALTAGRGKQKAKKKVLRFPASTSGVCMPLDEIRGTEVEVDQYGLA 264  
 QY 141 DTLITFDGELCDVFFY-----PAAQGE 164  
 Db 265 TLIVDDGKGLDDBHMQKTAFLLVGVITLALYKADDDGATLPSVDAMLADPKRDIG-E 323  
 QY 165 LRVYKSHYKAV 175  
 Db 324 LIMEEMATVGHV 334

RESULT 14		
DNAK_R1CPR		
ID	DNAK_R1CPR	STANDARD;
AC	Q9ZDY9;	PRT; 627 AA.
DT	30-MAY-2000 (Rel. 39, Created)	
DT	30-MAY-2000 (Rel. 39, Last sequence update)	
DT	28-FEB-2003 (Rel. 41, Last annotation update)	
DE	Chaperone protein dnak (Heat shock protein 70) (Heat shock 70 kDa protein) (HSP70).	
DE	GN	DNAK OR RP185.
OS	Rickettsia prowazekii.	
OC	Bacterialia, Proteobacteria, Alphaproteobacteria; Rickettsiales;	
OC	Rickettsiaceae; Rickettsiinae; Rickettsia.	
NCBI	taxid=782;	

[illegible]

this SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC  
DR EMBL; AJ235270; CAAL651.1; -.  
DR PIR; D17129; D11729.  
DR HSSP; P04475; IDc4.  
DR HAMAP; MF\_00332; -: 1.  
DR InterPro; IPR01023; Hsp70.  
DR Pfam; PF00012; HSF70; 1.  
DR PRINTS; PR00301; HEATSHOCK70.  
DR ProDom; PD00089; HSP70; 1.  
DR PROSITE; PS00287; HSF70\_1; 1.  
DR PROSITE; PS00289; HSF70\_2; 1.  
DR PROSITE; PS01036; HSP70\_3; 1.  
KM Chapterone; ATP-binding; Heat shock; Phosphorylation;  
KM complete proteome.  
FT MOD\_RES 197..197 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
SQ SEQUENCE 627 AA; 68383 MW; A658A0A55C8ACCT CRC64;

```

Query Match          7.2%; Score 81.5; DB 1; Length 627;
Best Local Similarity 28.2%; Pred. No. 14;
Matches    33; Conservative   15; Mismatches   42; Indels   27; Gaps    7

      72 GSQFCKLOYOE-----IERKEDYTVTSVFTEPNA---SSPIKYNTVETVKAVFG 121
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

```

Db      10 GTTNCVANNWEGKPKVIDAAGEGERTPTSIIFAFSESLVQGPAPKRAQVT----- 59
Qy     122 YKTNRNALAEVYGS--GLNTDTLLITDSELCDVYVYVYNNAGQGLAEVWKSKYHPV 176
Db      60 --NPNNTI-YAKKELIQNRNFTDPWAKDGLVY-ITVYKADNG--DAWVADNHNKTSIP 111

```

RESULT	ID	STANDARD;	PRT;	700 AA.
HS9C	DICDI			
AC	HS9C.DICDI			
AC	P54651;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Heat shock cognate 90 kDa protein.			
GN	HSPD OR HSC90.			
OS	Dicystosellum discoideum (Slime mold).			
OC	Dicystosella; Mycetozoa; Dictyostelidida; Dictyostelium.			
OX	NBRI_Taxid=44689;			

AN  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HX3;  
RA Boveré H., Dietrich W., Mintert U., Lottspeich F., Gerisch G.,  
PA Felix J.;  
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: Molecular chaperone. Has ATPase activity (By  
CC similarity).  
CC -1- SIMILARITY: Belongs to the heat shock protein 90 family.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

CC  
DR EMBL: U43591; AAA6917.1; -.  
DR HSSP: P07900; IYER.  
DR DickeyBase: DDB0001943; hepd.  
DR InterPro: IPR003594; ATPbind\_Atpase.  
DR InterPro: IPR001404; Hsp90.  
DR Pfam: PF02518; HATPase\_c/1.  
DR Pfam: PF00183; HSP90\_1.  
DR PRINTS: PR00775; HEATSHOCK90.  
DR SMART: SMO0387; HATPase\_c/1.  
DR PROSITE: PS00298; HSP90\_1.  
DR Chaperone; ATP-binding; heat shock.  
KW SEQUENCE 700 AA; 79881 MW; BFB5ECF238089CD3 CRC64;

Query Match 7.2%; Score 81; DB 1; Length 700;  
Best Local Similarity 19.5%; Pred. No. 17;  
Matches 52; Conservative 32; Mismatches 105; Indels 78; Gaps

```

QY 11 FVSALATQAETTSAAV--ENPLWA-----HEELLGK-----Y 42
Db 128 FYSALVALDPIVISHSKNDEEQYVWSSAGSEPTIALDHPPEPGRGRTKIVLHNKREDOLDY 18
QY 43 QDAWK-----SIDQGVSVTVLAKTYENDTDSWGSOPKC-LQVQEIERRKEDYT 91
Db 186 LDFIKINVLKXHSSEPIQYPSILILIKKEDVDEFTAKKEEESTDAKIEEIEERKKK 24
QY 92 V---TSVFTFRMSSPIKRYNTETVKAPOYGHKNRMLE-----YQCGGLNITD 14
Db 248 VKVQEKMDVIAKTKELTRNDSVDVTKKEZNSFYKXISNMHEEPLVKHFSVSGQLFRA 30

```

```

QY      191 -----RKYTDINE 199
Db      368 RETLOQNKLIVIRKLVKKCIETLENE 394
QY      308 ILVFKKAPFDLPESKKKANIKLYKSVFIMDNCADIIPYLINFLVRGIVDSBDPLNIS 367

```

Tue Mar 9 14:18:21 2004

us-09-555-296b-4.rsp

Page 10

Search completed: March 9, 2004, 13:34:32  
Job time : 20 secs

---



Db 121 GYKNNINALEYQVGGSLNTDILFTDDELCDVFFVFNADQCELMWKKSHYKAVPYCT 180  
 QY 181 FVFNVCADKRTYDIFNEECYNGEPWL 209  
 181 FVFNVCADKRTYDIFNEECYNGEPWL 209

## RESULT 2

Q8MW98 PRELIMINARY; PRT; 210 AA.  
 AC Q8MW98; 22, Created  
 DT 01-OCT-2002 (Tremblrel. 22, last sequence update)  
 DT 01-OCT-2002 (Tremblrel. 22, last annotation update)  
 DE Putative 22.5 kDa secreted protein.  
 OS Ixodes scapularis (Black-legged tick) (Deer tick).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodida; Ixodidae; Ixodes.  
 NC NCB1\_TaxID=6945;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Rhode Island; TISSUE=salivary gland;  
 RA Valenzuela J.G., Francischetti I.M., Pham V.M., Garfield M.,  
 RT "Exploring the Salivome of the Tick Vector of Lyme Disease, Ixodes  
 RT scapularis,"  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF483742; AAM93664.1; -  
 SQ SEQUENCE 210 AA; 24689 MW; C5B75614649A9BED CRC64;

Query Match 11.2%; Score 126; DB 5; Length 210;  
 Best Local Similarity 24.3%; Pred. No. 0.0029;  
 Matches 53; Conservative 41; Mismatches 90; Indels 34; Gaps 12;

QY 6 VLLITFVSALATQAEITTSKAGENPLMAHEELIGYQDAWKSIDQVSVTVYLAKTYYE 65  
 3 VMHPSFLCLLATIVD--AKPEIRI-DEDEKYMQYQIQALNPNPRESWLYYR-TYR 57  
 QY 66 NDQGSWGSQPKCL--QVOETERKEEDYTVSVTFPNASSPKYNTVETVAVVQYGYK 123  
 Db 58 REVD--GSEHICVSARKSENQNGSDYEVQVEYRLGT---KEQVTRKYV-TLYATPYK 109  
 QY 124 NINALEYQVGGSLNTD-----TLIFTDELCDVFFV--FNADQCELMW-KXS 170  
 110 TEMHATQRONNNMARVSQKDAQDGKYQLIYSDYKCDILVLENSGHDELYLSKA 169  
 QY 171 HYKAVPDYCTFVFNVCADKRTYDIFNEECYNGEPW 208  
 Db 170 LDDGVPRECEVYGIACGKDEPSY---KQRYV--YFW 201

## RESULT 3

Q8MW93 PRELIMINARY; PRT; 306 AA.  
 ID Q8MW93; 22, Created  
 DT 01-OCT-2002 (Tremblrel. 22, last sequence update)  
 DT 01-OCT-2002 (Tremblrel. 22, last annotation update)  
 DE Putative secreted histamine binding protein.  
 OS Ixodes scapularis (Black-legged tick) (Deer tick).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodida; Ixodidae; Ixodes.  
 NC NCB1\_TaxID=6945;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Rhode Island; TISSUE=salivary gland;  
 RA Valenzuela J.G., Francischetti I.M., Pham V.M., Garfield M.,  
 RT "Exploring the Salivome of the Tick Vector of Lyme Disease, Ixodes  
 RT scapularis,"  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF483717; AAM93639.1; -  
 SQ SEQUENCE 306 AA; 34232 MW; B2DB09181CC4395D CRC64;

Query Match 10.4%; Score 117.5; DB 5; Length 306;  
 Best Local Similarity 20.0%; Pred. No. 0.0027;  
 Matches 46; Conservative 34; Mismatches 73; Indels 77; Gaps 7;

QY 1 MCMQVLLITFVSALATQAEITTSKAGENPLMAHEELIGYQDAWKSIDQVSVTVYLA 60  
 1 MGLMYAALFACIASASAWQAOQITRNPNENL-LNEPSLPGSAMPAINKSKDPFLVM 59  
 Db 61 -----KTYENDTSGWSQPKQVOEIERKEDYT 91  
 QY 60 FSRMRHEINITCVVTASLHNETLKIYFRTYNNET---DGKNDLLEYQVRLNQDYK 116  
 Db 92 VTSV-----FTFNASSPKYNTVETVAVVQYGYKNT 125  
 QY 117 LENVTRAGIKGTPSDKPTPLGSNNYIEYGDYSCNTSSKPL---TMDLRA----- 163  
 QY 126 RMAIEYQVG-----GLNITDILFTDDELCDVFFVFNADQCELMWKKKS 170  
 Db 164 KDAVSGASASAPVEGVNIDPYYVHNPCHNIIRSLPKGCDPFLRK 213

## RESULT 4

Q819T9 PRELIMINARY; PRT; 176 AA.  
 ID Q819T9; 23, Created  
 DT 01-MAR-2003 (Tremblrel. 23, last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, last annotation update)  
 DT 01-JUN-2003 (Tremblrel. 24, last annotation update)  
 DE TSGP4.  
 OS Ornithodoros savignyi.  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodida; Argasidae; Ornithodoros.  
 NC NCB1\_TaxID=69826;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=salivary gland;  
 RX MEDLINE=21317941; PubMed=11425229;  
 RA Mans B.J., Venter J.D., Vrey P.J., Louw A.I., Netlz A.W.,  
 RT "Identification of putative proteins involved in granule biogenesis of  
 RT tick salivary glands,"  
 RL Electrophoresis 22:1739-1746 (2001).  
 DR EMBL; AF452891; AAN76831.1; -  
 SQ SEQUENCE 176 AA; 19375 MW; 44B7559F3BCBEA25 CRC64;

Query Match 9.9%; Score 111.5; DB 5; Length 176;  
 Best Local Similarity 19.2%; Pred. No. 0.044;  
 Matches 30; Conservative 33; Mismatches 64; Indels 29; Gaps 6;

QY 44 DAWKSIDQVSVTVYLAKTYYENDTSGWSQPKQVOEIERKEDYTVSVTFPNASS 103  
 Db 23 DVW-NVLKGSDEKFLMVYKRYTERGAN-----KCYMRKTSMDSSHTLEVMGSKAGT 75  
 QY 104 PIYK-----YNTETVKAVFOGYKNIR-----NAIEYQVGGSLNTDILFTDDELCDV 153  
 Db 76 TIDFVPSKTYTVAISSEGASTYNNMTVARGPSHGVKKE-----LYSDDGCGNI 125  
 QY 154 FVFNVA--DQCELMWKKSHYKHPDYCTFVFNVC 187  
 Db 126 LQKMTSPFGKCELMWFBGAKAVSSSGSKFKELC 161

## RESULT 5

Q8MW92 PRELIMINARY; PRT; 311 AA.  
 ID Q8MW92; 22, Created  
 DT 01-OCT-2002 (Tremblrel. 22, last sequence update)  
 DT 01-OCT-2002 (Tremblrel. 22, last annotation update)  
 DE Putative secreted histamine binding protein.  
 OS Ixodes scapularis (Black-legged tick) (Deer tick).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodida; Ixodidae; Ixodes.



OX NCB1\_TaxID=6945;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Node Island, TISSUE=salivary gland.  
 RA Valenzuela J.G., Francischetti I.M., Pham V.M., Garfield M.,  
 RA Macher T.N., Ribeiro J.M.C.,  
 RT "Exploring the Sialome of the Tick Vector of Lyme Disease, Ixodes  
 RT scapularis."  
 RT Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: AF483718; AAM93640.1;  
 DR EMBL: AF483718; AAM93640.1; D7B529FEF8ACFA9A CRCC4;  
 SQ SEQUENCE 311 AA; 36581 MW; D7B529FEF8ACFA9A CRCC4;

Query Match 9.2%; Score 103; DB 5; Length 311;  
 Best Local Similarity 19.6%; Pred. No. 0.54;  
 Matches 59; Conservative 35; Mismatches 79; Indels 128; Gaps 12;

QY 12 VSAALATQAEITTSKAGENPLMAHELLGKQDAWKSIDGVSVTYVLAKTYYENTGSM 71  
 DB 13 VSAAYAEVFOQMDRAPDNNPDLMKDLGAMQDAWRTIKFTANHSYLYSS-----GM 66  
 QY 72 GSQ-----FKCLQVQELERKEDYTVSVTFENASSPIKYNN-----VTEYKAVF 118  
 DB 67 GTRHYEDVRCIQVHSSDLN-----YTLKSANTSKYNNRTSKNNSSIQVYQAK 117  
 QY 119 Q--YGYKNIR-----NAIEYQV-----GGGLNT-----TDTLI 144  
 DB 118 QKYVSIENIMHGGPQREVTSPNGTCYNLNFLESGGCRHHOBQWQKWTYSEKTV 177  
 QY 145 FIDGELCDVFPVNPADQ-----CELM-----VKKSH----- 171  
 DB 178 LFTSTLCYVVSLODDSEYSCFNLSEDMIAKKNVTIPQVYTLLEKDSDEIKSEERSE 237  
 QY 172 -----YKAVPDYCTFVFNPCA-----KDKTYDIN 198  
 DB 238 SYRKESVQCEEPRTANKTTFLYDLFLFKELPSSCRAYFLNCGYPRYRKYDKDDKIN 297  
 QY 199 E 199  
 DB 298 E 298

## RESULT 6

OX 083750 PRELIMINARY; PRT; 722 AA.  
 AC 083750;  
 DT 01-JUN-2003 (Tremblrel. 24, Created)  
 DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
 DE Amino acid ABC transporter, amino acid-binding/permease protein.  
 GN EF0761.  
 OS Enterococcus faecalis (Streptococcus faecalis).  
 OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.  
 OX NCB1\_TaxID=1351;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=V583 / ATCC 700802;  
 RX MEDLINE=22550657; PubMed=12663927;  
 RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,  
 RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,  
 RA Tetelin H., Dodson R.J., Umayam L., Brinkac L., Beaman M.,  
 RA Daugherty S., Deboy R.T., Durkin S., Kolonay J., Madupu R., Nelson W.,  
 RA Vamathevan J., Tran B., Upton J., Hansen T., Shetty J., Khouli H.,  
 RA Utechback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.,  
 RT "Role of mobile DNA in the evolution of vancomycin-resistant  
 RT Enterococcus faecalis."  
 RL Science 293:2071-2074(2003).  
 DR EMBL: A8016949; AAO80578.1; -;  
 DR TIGR: EF0761;  
 DR GO: GO:0016020; C:membrane; IEA.  
 DR GO: GO:0030288; C:periplasmic space (sensu Gram-negative Bact. .; IEA.  
 DR GO: GO:0005334; F:glutamate-gated ion channel activity; IEA.  
 DR GO: GO:0004770; F:ionotropic glutamate receptor activity; IEA.  
 DR GO: GO:0005215; F:transporter activity; IEA.

DR GO: GO:0006810; P:transport; IEA.  
 DR Interpro: IPR000515; BPD\_transp.  
 DR Interpro: IPR001320; Ion\_glu\_receptor.  
 DR Interpro: IPR001311; SBP\_glu\_receptor.  
 DR Interpro: IPR001638; SBP\_bac\_3.  
 DR Pfam: PF00528; BPD\_transp; 1.  
 DR Pfam: PF00497; SBP\_bac\_3; 2.  
 DR SMART: SM00062; PBD; 2.  
 DR SMART: SM00079; PBPE; 1.  
 DR PROSITE: PS00402; BPD\_TRANSNP\_INN\_MEMBER; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 722 AA; 79713 MW; D7F6BB208DFE62 CRCC4;

Query Match 9.1%; Score 102.5; DB 16; Length 722;  
 Best Local Similarity 26.9%; Pred. No. 1.8;  
 Matches 60; Conservative 26; Mismatches 72; Indels 65; Gaps 15;

QY 16 LATQAEITSA-----KAGENPLMAHELLGKQDAWKSIDGVSVTYVLAKTYY----- 64  
 DB 204 LVGDKETGSSYGRAVKKQNP-----ELIKFPAAGLKLKNG---TYDKILNNTYATGD 255  
 QY 65 ENDTGSWGSQFKCLQVQELERKEDYTVSVTFENASSPIKYNN-----TETVKA 116  
 DB 256 ETTQDAGEQMK-----KITPKKEKYIASDSTF---APFEQNAQGDYVIGIDVLR 306  
 QY 117 V-----FQY-GYKNIRNALEY-QVGG---GLNITDTLFTDGLCDVFPVNPADQ 162  
 DB 307 AAELQGTVEKFTIGSSAVQAVESGQADQMTATITD-----DKKAPDFEVPYFDSG 361  
 QY 163 CELWKKSH--YKAVPDYCTFVFN-----FCAKDKTYD 195  
 DB 362 IQIAVKKGNDKIKSYDDLKGGKYGVKIGTESADFLKNNKKYD 404

## RESULT 7

ID Q95W25 PRELIMINARY; PRT; 221 AA.  
 AC Q95W25;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)  
 DE Histamine binding protein.  
 GN HBP.  
 OS Ixodes scapularis (Black-legged tick) (Deer tick).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodida; Ixodidae; Ixodes.  
 OX NCB1\_TaxID=6945;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Das S., Banerjee G., Deponte K., Marcantonio N., Kantor F.S.,  
 RA Fikrig E.,  
 RT "Salp2SD, an Ixodes scapularis antioxidant, is one of 14  
 RT immunodominant antigens in engorged tick salivary glands."  
 RL J. Infect. Dis. 184:10-0(2001).  
 DR EMBL: AF209913; AAK97816.1; -;  
 DR EMBL: AF209913; AAK97816.1; 7B95ACGCAFA252 CRCC4;  
 SQ SEQUENCE 221 AA; 25359 MW; 7B95ACGCAFA252 CRCC4;

Query Match 8.7%; Score 97.5; DB 5; Length 221;  
 Best Local Similarity 23.4%; Pred. No. 1.1;  
 Matches 49; Conservative 43; Mismatches 80; Indels 37; Gaps 14;

QY 10 TFVSAALATQAEITTSKAGENPLMAHELLGKQDAWKSIDGVSVTYVLA 60  
 DB 31 TVGSIGTTRRGATGARYVTAPPEEDP-----SKYKQVATVVE--MNAQWQW 80  
 QY 61 KTTYENDTGSWGSQFKCLQVQELERKEDYTVSVTFENASSPIKYNNVET--VKAV 117  
 DB 81 WRITD-VTDPSGNTVQCEPNVMEKR---TPTNYSQVYKSKN-SWETIDETLLIKDI 134  
 QY 118 FOYGYKNIRNALEYQVGGGLNITDTL-FTDGLCDVFPVNPADQ---CELWKK-SHY 172  
 DB 135 GHHGPPNVMNFORPTIG---IATNLTLYSNVNCVLRIPFTVQGRHCDLMMANLTL 191

Qy 173 KHWDPYCTFVFNVCADKDKTYDIFNEEC 201  
 Db 192 QETPDDCLNKPFVYC-NTTQIYRVYVPS 219

## RESULT 8

Q8A286 PRELIMINARY; PRT; 508 AA.  
 AC Q8A286;  
 DT 01-JUN-2003 (Tremblrel, 24, Created)  
 DT 01-JUN-2003 (Tremblrel, 24, Last sequence update)  
 DT 01-JUN-2003 (Tremblrel, 24, Last annotation update)  
 DE Hypothetical protein.  
 GN BT3420.  
 OS Bacteroides thetaiotaomicron.  
 OC Bacteroidetes; Bacteroidetes; Bacteroides (class); Bacteroidales;  
 OC Bacteroidaceae; Bacteroides.  
 CX NCBI\_TaxID=818;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VPI-5482 / ATCC 29148;  
 RX MEDLINE=22550858; PubMed=12663928;  
 RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,  
 RA Chang H.C., Hooper L.V., Gordon J.I.;  
 RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis."  
 RL Science 299:2074-2076(2003).  
 DR EMBL; AF016940; AAC078526.1;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 508 AA; 59877 MW; B80E9E38262d6450 CRC64;

## Query Match

Best Local Similarity 8.4%; Score 95; DB 16; Length 508;  
 Matches 43; Conservative 27; Mismatches 64; Indels 62; Gaps 9;

Qy 6 VLLITFVSALATQ-----AETSAKAGENLMAHEILGKYODAWKSID-----QGV 54  
 Db 8 IVSLFFPYTGDDTQCTKPPSVETVLAAGNN-----RKELEALDLYFKQGS 56  
 Qy 55 VT-----YLAKTTYNDSGSGQPKCQVEIRKEDVTYVSVFTRASSPIKTY 108  
 Db 57 IKIKAIFFLVAMHDINHSEITYWKTR-----GKRVDF-----SEFDTA 95  
 Qy 109 NTEETKAVFOYGYKYNRAIEYOVGGSLNTDTLFTTGELCDVFPVNAOQCELWVK 168  
 Db 96 NLETKAVKALDSMRKY-----GSLDFQDTIIVDYNLSLGLKVLNNVNAVDTW-R 144  
 Qy 169 KSHYKVP--DYCTFV 182  
 Db 145 LSEYKDIPFNDCEVI 160

## RESULT 9

Q8IDR7 PRELIMINARY; PRT; 972 AA.  
 AC Q8IDR7;  
 DT 01-MAR-2003 (Tremblrel, 23, Created)  
 DT 01-MAR-2003 (Tremblrel, 23, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel, 23, Last annotation update)  
 DE Hypothetical protein.  
 GN PF13\_0230.  
 OS Plasmodium falciparum (isolate 3D7).  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 CX NCBI\_TaxID=36329;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Harris B., Lennard N., Clark L., Lane A., Barron A., Corton C.,  
 RA Berrieman M., Bain A., Hall N., Atkin R., Chillingworth C., Doggett J.,  
 RA Ormond D., Sanders W., Hayes R., Hall S., Quail M., Barrett B.,  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL844509; CAD52552.1;  
 KW Hypothetical protein.  
 SQ SEQUENCE 972 AA; 119242 MW; 70471A60C0327C32 CRC64;

Query Match 7.9%; Score 89; DB 5; Length 972;  
 Best Local Similarity 19.8%; Pred. No. 43;  
 Matches 43; Conservative 33; Mismatches 81; Indels 60; Gaps 9;

Qy 4 QVLLITFVSALATQAEETSAKAGENLMAHEILGKYODAWKSIDGV-----SVTV 58  
 Db 289 KICLFPIYISCSIVSQTYKHMKNN-----VEKKKKSLDFFIKINECKHI 340  
 Qy 59 LAKTYENDTSGSGQPKCQVEIE-----RKEDYTVSVTFENASSPIKYNTVETV 114  
 Db 341 LMDNLN-----KINDQYIRYIIRKRCSTYANTVLAQKYNLSIKYINIIINIK 390  
 Qy 115 KAVFQ--YGYK---IRNAIEYOVGGSLNTDTLFTTGELCD--VEFPVNAOQCELWVK 168  
 Db 391 KSPFYCYLYNSYFNKSIY-----CSLFIYD-----LOKY 423  
 Qy 169 KSHYKVPDYCTFVFNVCADKDKTYDIFNEECVYNG 205  
 Db 424 NLFYTYLQILLIKFNLFLDKRKTYNDYKHIILYG 460

## RESULT 10

Q939L2 PRELIMINARY; PRT; 598 AA.  
 AC Q939L2;  
 DT 01-DEC-2001 (Tremblrel, 19, Created)  
 DT 01-DEC-2001 (Tremblrel, 19, Last sequence update)  
 DT 01-JUN-2003 (Tremblrel, 24, Last annotation update)  
 DE Arylsulfate sulfoesterase.  
 GN ASTA.  
 OS Citrobacter freundii.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Citrobacter.  
 CX NCBI\_TaxID=546;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MB4-8242;  
 RX MEDLINE=21425731; PubMed=11534764;  
 RA Kang J.W., Jeong Y.J., Kwon A.R., Yun H.J., Kim D.H., Choi E.C.;  
 RT "Cloning, sequence analysis, and characterization of the astA gene  
 encoding an arylsulfate sulfoesterase from Citrobacter freundii."  
 RL Arch. Pharm. Res. 24:316-322(2001).  
 DR EMBL; AY029222; AAK31641.1;  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 KW Transferase.  
 SQ SEQUENCE 598 AA; 66501 MW; 4CE9BB26545F62A5 CRC64;

## Query Match

Best Local Similarity 7.7%; Score 87; DB 2; Length 598;  
 Matches 46; Conservative 25; Mismatches 55; Indels 66; Gaps 10;

Qy 27 AGENPLMAHE-----LIGKYODAWKSIDQSVTVYLAQTYEN----- 66  
 Db 377 AGEN--MAVNSIAYANDSITLSSRHQGVK-IGRKQYKWLAPKGNKALAKSL 433  
 Qy 67 -DTSGSGQPKCQVEIRKEDYTV-----SVTFER-----ASSPIK 106  
 Db 434 KPYDDKGNALCKDENKCENTDFTYDHTAMSSKGLTITFDGDRGLBQPLPTWK 493  
 Qy 107 Y-----YNTE---TYKAVFOYG-----YKINRAIEYQ-----VGGGLNTDTL 143  
 Db 494 YSFYKIKDKKGVYQVWYVYKGRGDFPISVLEYQDRPTMGFGGSINLPVVG 553  
 Qy 144 IFTDGLCDVY 155  
 Db 554 OPTIKINEIDY 565

## RESULT 11

Q92N47 PRELIMINARY; PRT; 932 AA.  
 AC Q92N47;  
 DT 01-DEC-2001 (Tremblrel, 19, Created)

DT 01-DEC-2001 (Tremblrel. 19, last sequence update)  
 DT 01-JUN-2003 (Tremblrel. 24, last annotation update)  
 DE Putative outer membrane receptor protein.  
 GN R02833 OR SMC02721.  
 OS Rhizobium meliloti (Sinorhizobium meliloti).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.  
 OX NCBI\_TaxID=382;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1021;  
 RX MEDLINE=21396507; PubMed=11481430;  
 RA Capela D., Barloy-Hubler F., Gonzy J., Bothe G., Ampe F., Batut J.,  
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,  
 RA Pohl T., Portetelle D., Puehler A., Purnelle B., Rampeger U.,  
 RA Renard C., Thebaud P., Vanderbol M., Weidner S., Galibert F.,  
 RT "Analysis of the chromosome sequence of the legume symbiont  
 RT Sinorhizobium meliloti strain 1021."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
 RL EMBL; AL591790; CAC46962.1; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0005215; F:transporter activity; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR000531; T0NB\_BoxC.  
 DR Pfam; PF00593; T0NB\_dep\_Rec; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 932 AA; 101872 MW; E4F6473BAF5AC484 CRC64;

Query Match 7.7%; Score 86.5; DB 16; Length 932;  
 Best Local Similarity 25.5%; Pred. No. 68;  
 Matches 41; Conservative 24; Mismatches 47; Indels 49; Gaps 11;

QY 44 DAMKSIDQSVTVYIAKTYENDTG---SMGSG--FKCLQ--QVEIERKEEDYTVSVF 96  
 DB 221 DDFKXIDRG-----GDTGFDLTDWGGKLGFRADGTQNEPFKXQNYDETSNE 267  
 QY 97 TFRNMS-----SPKRYN-----VTETKAVFQYGYKYNINAEYQVGGGLNTDTLIF 145  
 DB 268 TYLGTLTDFKDTPRRRYNASQKVIDAANHNYQFSH-----SYEITPDINLTITVYR 320  
 QY 146 TDG-----ELCDVFYVPNADQCELMVKKSHYKVPD-YCT 180  
 DB 321 TDTTRAWYKLDNV--RNDSDTG--WVSLSNILADPDYIST 356

RESULT 12  
 Q8XMD0 PRELIMINARY; PRT; 399 AA.  
 AC Q8XMD0;  
 DT 01-MAR-2002 (Tremblrel. 20, Created)  
 DT 01-MAR-2002 (Tremblrel. 20, last sequence update)  
 DT 01-JUN-2003 (Tremblrel. 24, last annotation update)  
 DE Hypothetical protein CPE0759.  
 GN CPE0759.  
 OS Clostridium perfringens.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 OX NCBI\_TaxID=1502;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=13 / Type A;  
 RX MEDLINE=21664373; PubMed=11792842;  
 RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,  
 RA Shibata T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.,  
 RT "Complete genome sequence of Clostridium perfringens, an anaerobic  
 RT flesh-eater."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).  
 DR EMBL; AP003189; BAB80465.1; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 399 AA; 46802 MW; 61019F007582B815 CRC64;

Query Match 7.6%; Score 86; DB 16; Length 399;  
 Best Local Similarity 23.8%; Pred. No. 24;  
 Matches 30; Conservative 21; Mismatches 37; Indels 38; Gaps 4;

QY 42 YQDAKS---IDG-----VSVTVIAKTYENDTGSGSQPKCLQVQEIET 85  
 DB 224 YEEAKSLIEYIESGFNLMCSRSENIMNLFSRIRRTSNIRKASYEAKFALSMIDM-- 281  
 QY 86 KEBDYTVSVTFPNASPIKYV-----TETKAVFQYGYKYNINAEYQVGGGL 137  
 DB 282 -----YNDSKILIKNELGIIRLFIEIAIKVDYQCYENIGIILBDSHGM 329  
 QY 138 NITDTL 143  
 DB 330 NLEETL 335

RESULT 13  
 Q51717 PRELIMINARY; PRT; 584 AA.  
 AC Q51717;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25, last annotation update)  
 DE Hypothetical protein precursor.  
 OS Pseudomonas fluorescens.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=294;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DSM 7139;  
 RX MEDLINE=95050280; PubMed=7961472;  
 RA Schlimmer A., Jendrossek D.,  
 RT "Molecular characterization of the extracellular poly(3-  
 RT hydroxyoctanoic acid) [P(3HO)] depolymerase gene of Pseudomonas  
 RT fluorescens GKL3 and of its gene product."  
 RL J. Bacteriol. 176:7065-7073(1994).  
 DR EMBL; U10470; AAA64539.1; -.  
 KW Hypothetical protein; Signal.  
 FT SIGNAL 1 26  
 SQ SEQUENCE 584 AA; 63249 MW; 3A16FA898D2CBF CRC64;

Query Match 7.6%; Score 86; DB 2; Length 584;  
 Best Local Similarity 27.1%; Pred. No. 40;  
 Matches 29; Conservative 15; Mismatches 51; Indels 12; Gaps 2;

QY 39 LGRYQDAKSIDQSVTVYIAKTYENDTGSGSQPKCL-----QVQEIETKEEDYT 91  
 DB 433 VQTHLDGYREKDM-----YQASLAIFYNFGVAGADATLFTLVASRVQGLSDVDYYE 487  
 QY 92 VTSVTFPNASPIKYVNTETKAVFQYGYKYNINAEYQVGGGL 138  
 DB 488 ATSTATGAQASLQLYTNVNLVNLVPSFQYQISINCVAQLTNGLD 534

RESULT 14  
 Q33983 PRELIMINARY; PRT; 917 AA.  
 AC Q33983;  
 DT 01-JAN-1998 (Tremblrel. 05, Created)  
 DT 01-JAN-1998 (Tremblrel. 05, last sequence update)  
 DT 01-JUN-2003 (Tremblrel. 24, last annotation update)  
 DE AMI.  
 GN AMI.  
 OS Listeria monocytogenes.  
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.  
 OX NCBI\_TaxID=1639;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=EGD;  
 RX MEDLINE=97426036; PubMed=9282740;  
 RA Braun L., Dramsi S., Dehoux P., Bierre H., Lindhal G., Cossart P.;

RT "InlB : an invasion protein of *Listeria monocytogenes* with a novel  
RT type of surface association.";  
RL Mol. Microbiol. 25:285-294(1997).  
DR EMBL; U82488; AAC45605.1; -  
DR GO; GO:0008745; F:N-acetylmuramoyl-L-alanine amidase activity; IEA.  
DR GO; GO:0009253; P:peptidoglycan catabolism; IEA.  
DR InterPro; IPR002502; Amidase\_2.  
DR Pfam; PF01510; Amidase\_2; 1.  
DR SMART; SM00644; Ami\_2; 1.  
SQ SEQUENCE 917 AA; 102352 MW; 21EA2452B4F103B7 CRC64;

Query Match 7.6%; Score 85.5; DB 2; Length 917;  
Best Local Similarity 23.2%; Pred. No. 81;  
Matches 38; Conservative 24; Mismatches 49; Indels 53; Gaps 9;

QY 28 GENPLMAHSELKIQDAMKSIDQSVTVYLAITYENDT---GSGW-----S 73  
DB 223 GKGTWMSH-----AVTRYLGITHTDPVAFNQMGNFNNFVSLINE 265  
QY 74 QFKCLQYQEI ERKEDYTVTVTFNASSPIKY--NVTETKAV--FOYGENIR-- 126  
DB 266 KYKAMQY-NYEKEIYDQALTAYSRVKATGNSVWTKPKTEGAKLVNPLSSYSGKRLRI 324  
QY 127 -----NAIEYQ--VGGGLNTDTLFTDGLCDVYFVNADQ 161  
DB 325 REAKTSGGTWYQFSVVG-----KITGWDSKALNTFTPSMEK 363

## RESULT 15

Q8A812 PRELIMINARY; PRT; 1090 AA.  
ID Q8A812  
AC Q8A812;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE OmpA-related protein.  
GN Btl185.  
OS Bacteroides thetaiotaomicron.  
OC Bacteria; Bacteroidetes; Bacteroidales;  
OC Bacteroidaceae; Bacteroides.  
OX NCBI\_TaxID=818;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VPI-5482 / ATCC 29148;  
RX MEDLINE=22550858; PubMed=1263928;  
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,  
RA Chang H.C., Hooper L.V., Gordon J.I.;  
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis";  
Science 299:2074-2076(2003).  
RL EMBL; AEO16930; AAC76292.1; -  
DR GO; GO:0004553; P:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.  
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
DR InterPro; IPR008969; Carboxypep\_reg.  
DR InterPro; IPR001064; Crystallin.  
DR InterPro; IPR001111; Glyco\_hydro\_GHD.  
DR ProDom; PD002572; Glyco\_hydro\_GHD; 1.  
DR PROSITE; PS00225; CRYSTALLIN\_BETAGAMMA; 1.  
KW Complete proteome.  
SQ SEQUENCE 1090 AA; 122044 MW; 63C8B3A1B8A2DFN6 CRC64;

Query Match 7.6%; Score 85.5; DB 16; Length 1090;  
Best Local Similarity 23.8%; Pred. No. 1e-02;  
Matches 44; Conservative 19; Mismatches 81; Indels 41; Gaps 9;

QY 47 KSIDQ-----VSTVYLAITYEN---DTGSGSQFKLQYQEI ERKEDYTVTV--F 96  
DB 821 KKYDFGTGSGIGYTTAKDLTANPGSAPNSAMQNN---VAVNSLNDPGVSYSLFSTPHR 877  
QY 97 TFRNASSPIKTYNTEVKAVFQYKYNIRNALEY--QVGGGLNTDTLFTDGLCDVF 154  
DB 878 IIAVASYEINVAKLKLTFLFSGVHTGRSYSTYYNDMG-----DGNYSDLI 926  
QY 155 YVFNADQCEL-----WVKSHTKHPVDYCTFVFNFCADKRTYDIFNEECVNGE 206

DB 927 YVFNADQELTVDITDKSGAITYSVVDQAKDFWDFVNDSTLKDRK-----GKYVERNGS 981  
QY 207 --PWL 209  
DB 982 LTPWI 986  
Search completed: March 9, 2004, 13:35:31  
Job time : 50 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 9, 2004, 13:30:30 ; Search time 60 seconds  
(without alignments)  
984.207 Million cell updates/sec

Title: US-09-555-296b-4

Sequence: 1 WKQVVLITFVSAALATQA.....DRKTYDIFNECVNGEPMWL 209

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23jan04:\*

1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1125	100.0	209	AAW37449	AAW37449 Tick Vaso
2	1125	100.0	209	AAW37449	AAW37449 Tick Vaso
3	364	32.4	203	AAW18081	AAW18081 Histamine
4	353	31.4	203	AAW18085	AAW18085 Histamine
5	290.5	25.8	190	AAW37447	AAW37447 Tick Vaso
6	290.5	25.8	190	AAW18079	AAW18079 Histamine
7	284.5	25.3	198	AAW18084	AAW18084 Histamine
8	282	25.1	171	AAW73261	AAW73261 Histamine
9	282	25.1	171	AAW74289	AAW74289 Histamine
10	265	23.6	200	AAW37448	AAW37448 Tick Vaso
11	265	23.6	200	AAW18080	AAW18080 Histamine
12	252	22.4	190	AAW37446	AAW37446 Tick Vaso
13	252	22.4	190	AAW18078	AAW18078 Histamine
14	251.5	22.4	182	AAW73262	AAW73262 Histamine
15	239	21.2	172	AAW73260	AAW73260 Histamine
16	239	21.2	172	AAW74288	AAW74288 Histamine
17	140	12.4	284	AAW18088	AAW18088 Histamine
18	127.5	11.3	285	AAW18087	AAW18087 Histamine
19	126.5	11.2	321	AAW18089	AAW18089 Histamine
20	102.5	9.1	722	AAW35203	AAW35203 Enterococ
21	97.5	8.7	722	AAW37764	AAW37764 I. scapu
22	96	8.5	650	AAW19265	AAW19265 Protein e
23	91	8.1	202	AAW37467	AAW37467 Sequence
24	88	7.8	644	AAW3830	AAW3830 Protein e

## ALIGNMENTS

26	87	7.7	639	6	ABU48728	ABU48728 Protein e
27	86	7.6	2778	4	ABBS5863	ABBS5863 Drosophi
28	85	7.6	254	4	AAW17991	AAW17991 Arabidops
29	85	7.6	411	4	ABG03896	ABG03896 Novel hum
30	85	7.6	437	3	AAW17990	AAW17990 Arabidops
31	85	7.6	441	3	AAW17989	AAW17989 Arabidops
32	85	7.6	656	6	ABU35973	ABU35973 Protein e
33	84.5	7.5	485	6	ABU29629	ABU29629 Protein e
34	84.5	7.5	497	7	ADCG95652	ADCG95652 E. faecit
35	84.5	7.5	917	5	ABW48437	ABW48437 Listeria
36	84	7.5	217	5	AAW97773	AAW97773 I. scapu
37	83	7.4	207	2	AAW18083	AAW18083 Histamine
38	83	7.4	2285	2	AAW98149	AAW98149 Bacillus
39	82	7.3	555	4	ABW66457	ABW66457 Drosophi
40	82	7.3	876	6	ABW66456	ABW66456 Drosophi
41	81.5	7.2	775	6	ABU20485	ABU20485 Protein e
42	81	7.2	659	2	AAW24123	AAW24123 Thermococ
43	81	7.2	659	2	AAW24121	AAW24121 Thermococ
44	81	7.2	659	2	AAW94840	AAW94840 M0985925
45	80.5	7.2	597	2	AAW20536	AAW20536 H. pylori

RESULT 1  
ID AAW37449  
ID AAW37449 standard; protein, 209 AA.

AC AAW37449;  
XX  
XX 27-NOV-1997.  
DT 08-NOV-1998 (first entry)

XX Tick vasoactive amine binding protein D.RET6.

XX Vasoactive amine binding protein; D.RET6; histamine; serotonin; assay;  
KW antihistamine; anti-inflammatory; insect bite; snake bite; scorpion bite;  
KW dermatitis; vaccine; transgenic animal; tick.

XX Dermacentor reticulatus.

XX Key  
XX Peptide  
XX Location/Qualifiers  
XX 1..28  
XX /label= sig\_peptide

XX WO974451-A2.

XX 27-NOV-1997.

XX 19-MAY-1997; 97WO-GB001372.

XX 18-MAY-1996; 96GB-00010484.

XX 18-APR-1997; 97GB-00007844.

XX (OXFO-) OXFORD VACS LTD.

XX Paesen GC, Nuttall PA;

XX WPI, 1998-018506/02.

XX N-PSDB; AAW00230.

XX New vasoactive amine binding proteins and related nucleic acid, vectors -  
XX transformed cells and transgenic animals, used for assaying or removing  
XX histamine and as antihistamine or anti-inflammatory agents.

XX Example 2; Fig 4; 44pp; English.

CC This protein comprises tick Dermacentor reticulatus (Dr) novel vasoactive  
CC amine binding protein (VABP) D.RET6. Its amino acid sequence was deduced  
CC from a cDNA clone (see AAV00230) obtained from a Dr salivary gland cDNA  
CC library. 3 Novel VASPs, designated FS-HBP1, FS-HBP2 and MS-HBP1 (see  
CC AAW37446-48), of the tick Rhipicephalus appendiculatus have also been

CC identified. The VABPs can be expressed in host cells using e.g. a  
 CC baculovirus expression system. They can be used: (i) to assay histamine  
 CC (or other VA such as serotonin) in body fluids or cell culture  
 CC supernatants; e.g. to monitor the effect of allergens; (ii) for binding  
 CC VA, e.g. to remove histamine from blood, food, cell cultures etc.; (iii)  
 CC as an antihistamine or anti-inflammatory agents; e.g. for treating  
 CC insect, snake or scorpion bites or dermatitis, or as a carrier for slow  
 CC release of histamine-related compounds; (iv) in vaccines to protect  
 CC against metazoan parasites, especially in animals; (v) as reagents for  
 CC studying inflammation, involvement of VA in ulcer formation or the immune  
 CC response etc. VABPs provide a more sensitive assay for histamine than low  
 CC affinity antibodies currently used. They may also be more effective and  
 CC safer than conventional antihistamines. (Updated on 27-AUG-2003 to  
 CC correct OS field.)  
 XX  
 SQ Sequence 209 AA;

Query Match 100.0%; Score 1125; DB 2; Length 209;  
 Best Local Similarity 100.0%; Pred. No. 6.9e-104;  
 Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKQVLLITFVSALATQAEFTTSKAGENPLMAHEELLGKYQDAMKSIDQGVSTVYLA 60  
 DB 1 MKQVLLITFVSALATQAEFTTSKAGENPLMAHEELLGKYQDAMKSIDQGVSTVYLA 60  
 QY 61 KTTYENDTGSWGSQFCKLQVQEI ERKEEDYTVSVTFPRNASPIKYVNTETVAVAFQY 120  
 DB 61 KTTYENDTGSWGSQFCKLQVQEI ERKEEDYTVSVTFPRNASPIKYVNTETVAVAFQY 120  
 QY 121 GYKNIRNAIEYQVGGGLNTDTLLIFTDGELCDVFYVPNADQCELMVKSKHYKAVPYCT 180  
 DB 121 GYKNIRNAIEYQVGGGLNTDTLLIFTDGELCDVFYVPNADQCELMVKSKHYKAVPYCT 180  
 QY 181 FVFNVCACDKRKYTDIFNEECVYNGEPM 209  
 DB 181 FVFNVCACDKRKYTDIFNEECVYNGEPM 209

RESULT 2  
 ID AAY18081 standard; protein; 209 AA.

AC AAY18081;

DT 06-AUG-1999 (first entry)

DE Histamine binding protein D.RETG.

XX Histamine binding protein; serotonin binding compound; inflammation;  
 KW gastric acid secretion; allergy; type I hypersensitivity reaction;  
 KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;  
 KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;  
 KW respiratory disease; coronary heart disease; cellular growth regulator;  
 KW tissue repair; blood-sucking ectoparasite; therapy.

OS Rhipicephalus appendiculatus.

XX WO9927104-A1.

XX 03-JUN-1999.

XX 26-NOV-1998; 98WO-GB003530.

XX 26-NOV-1997; 97GB-00025046.

XX 26-JUN-1998; 98GB-00013917.

XX (OXFO-) OXFORD VACS LTD.

XX Nuttall PA, Paesen GC;

XX WPI; 1999-357841/30.

XX N-PSDB; AAX76967.

FT Histamine and serotonin binding compounds useful for the treatment of  
 PS allergies.  
 XX  
 XX Claim 1; Fig 4; 84pp; English.

XX This sequence is an example of a histamine or serotonin binding compound  
 CC (A) of the invention. The compounds are useful for regulating the action  
 CC of histamine and serotonin (in e.g. inflammation and gastric acid  
 CC secretion), the detection, quantification and removal of histamine or  
 CC serotonin (in animals, plants, cell cultures, food materials, or humans)  
 CC and in the treatment of various diseases and allergies (e.g. type I  
 CC hypersensitivity reactions, urticaria, asthma, allergic rhinitis (hay  
 CC fever), atopic dermatitis, insect bites and food and drug allergies,  
 CC abnormal blood pressure, migraine, psychological disorders, respiratory  
 CC disease, and coronary heart disease). Histamine may also be used to  
 CC regulate cellular growth and tissue repair. The molecules may also be  
 CC used as components of vaccines directed against blood-sucking  
 CC ectoparasites

SQ Sequence 209 AA;

Query Match 100.0%; Score 1125; DB 2; Length 209;  
 Best Local Similarity 100.0%; Pred. No. 6.9e-104;  
 Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKQVLLITFVSALATQAEFTTSKAGENPLMAHEELLGKYQDAMKSIDQGVSTVYLA 60  
 DB 1 MKQVLLITFVSALATQAEFTTSKAGENPLMAHEELLGKYQDAMKSIDQGVSTVYLA 60  
 QY 61 KTTYENDTGSWGSQFCKLQVQEI ERKEEDYTVSVTFPRNASPIKYVNTETVAVAFQY 120  
 DB 61 KTTYENDTGSWGSQFCKLQVQEI ERKEEDYTVSVTFPRNASPIKYVNTETVAVAFQY 120  
 QY 121 GYKNIRNAIEYQVGGGLNTDTLLIFTDGELCDVFYVPNADQCELMVKSKHYKAVPYCT 180  
 DB 121 GYKNIRNAIEYQVGGGLNTDTLLIFTDGELCDVFYVPNADQCELMVKSKHYKAVPYCT 180  
 QY 181 FVFNVCACDKRKYTDIFNEECVYNGEPM 209  
 DB 181 FVFNVCACDKRKYTDIFNEECVYNGEPM 209

RESULT 3

ID AAY18086 standard; protein; 203 AA.

AC AAY18086;

DT 06-AUG-1999 (first entry)

DE Histamine binding protein Ih/Bm-HBP2.

XX Histamine binding protein; serotonin binding compound; inflammation;  
 KW gastric acid secretion; allergy; type I hypersensitivity reaction;  
 KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;  
 KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;  
 KW respiratory disease; coronary heart disease; cellular growth regulator;  
 KW tissue repair; blood-sucking ectoparasite; therapy.

XX Boophilus microplus.

XX Ixodes hexagonus.

XX Synthetic.

XX WO9927104-A1.

XX 03-JUN-1999.

XX 26-NOV-1998; 98WO-GB003530.

XX 26-NOV-1997; 97GB-00025046.

XX 26-JUN-1998; 98GB-00013917.

XX (OXFO-) OXFORD VACS LTD.

XX Nuttall PA, Paesen GC,  
 PI WPI; 1999-357841/30.  
 DR N-PSDB; AAX76971.  
 XX  
 PT Histamine and serotonin binding compounds useful for the treatment of  
 PT allergies.  
 XX  
 PS Claim 14; Fig 8; 84pp; English.  
 CC This sequence is an example of a histamine or serotonin binding compound  
 CC (A), of the invention. cDNA encoding this sequence was isolated from a  
 CC mixed *Boophilus microplus*/Ixodes hexagonus cDNA expression library. The  
 CC compounds are useful for regulating the action of histamine and serotonin  
 CC (in e.g. inflammation and gastric acid secretion), the detection,  
 CC quantification and removal of histamine or serotonin (in animals, plants,  
 CC cell cultures, food materials, or humans) and in the treatment of various  
 CC diseases and allergies (e.g. type I hypersensitivity reactions,  
 CC urticaria, asthma, allergic rhinitis (hay fever), atopic dermatitis,  
 CC insect bites and food and drug allergies, abnormal blood pressure,  
 CC migraine, psychological disorders, respiratory disease, and coronary  
 CC heart disease). Histamine may also be used to regulate cellular growth  
 CC and tissue repair. The molecules may also be used as components of  
 CC vaccines directed against blood-sucking ectoparasites  
 XX  
 SQ Sequence 203 AA;  
 Query Match 32.4%; Score 364; DB 2; Length 203;  
 Best Local Similarity 38.1%; Pred. No. 6.3e-28;  
 Matches 75; Conservative 38; Mismatches 78; Indels 6; Gaps 5;  
 Db 6 VLLTFVSAALATQAEITTSKAGENPLMAHEELIKGYODAKWSIDGVSATYLAATY 65  
 8 VVYLTAVTAADAPSSSTRNEPLEKTTM-HNQTGRQDAMKSNQGVGTYTFLNSTYN 66  
 Qy 66 NDTGWSGQFKCLQVOEIERKEEDYTVTSVTFPNASSPIKYNTVETVAKVFOYGYKNI 125  
 Db 67 NDS-VWGNKFTCLSVTSKYESTFTVEYNTTYKNQSQ--QWVSMSENVTAVOEGGY-SV 122  
 Qy 126 RAIRFYOGGGLNITDPLTFDGLCDVFYVPNADGCELMVKSHYKHPDYCTFPFNV 185  
 Db 123 KNIQWTTENTKFNQDVIYFTDQCTDVLIFPYKEDGYELMVRSEYIQNTPTCCQPIFDL 182  
 Qy 186 FCADKRTYDIFNEECV 202  
 Db 183 -VALGRTYNISTPNCV 198  
 RESULT 4  
 AAY18085  
 ID AAY18085 standard; protein; 203 AA.  
 XX  
 AC AAY18085;  
 XX  
 DT 06-JUN-1999 (first entry)  
 XX  
 DE Histamine binding protein 1h/Bm-HBPI.  
 XX  
 KW Histamine binding protein; serotonin binding compound; inflammation;  
 KW gastric acid secretion; allergy; type I hypersensitivity reaction;  
 KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;  
 KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;  
 KW respiratory disease; coronary heart disease; cellular growth regulator;  
 KW tissue repair; blood-sucking ectoparasite; therapy.  
 XX  
 OS *Boophilus microplus*.  
 OS Ixodes hexagonus.  
 OS Synthetic.  
 XX  
 PN WO9927104-A1.  
 XX  
 PD 03-JUN-1999.

XX  
 PF 26-NOV-1998; 98WO-GB003530.  
 XX  
 PR 26-NOV-1997; 97GB-00025046.  
 PR 26-JUN-1998; 98GB-00013917.  
 XX  
 PA (OXFO-) OXFORD VACS LTD.  
 XX  
 XX Nuttall PA, Paesen GC,  
 PI WPI; 1999-357841/30.  
 DR N-PSDB; AAX76970.  
 XX  
 PT Histamine and serotonin binding compounds useful for the treatment of  
 PT allergies.  
 XX  
 PS Claim 13; Fig 7; 84pp; English.  
 CC This sequence is an example of a histamine or serotonin binding compound  
 CC (A), of the invention. cDNA encoding this sequence was isolated from a  
 CC mixed *Boophilus microplus*/Ixodes hexagonus cDNA expression library. The  
 CC compounds are useful for regulating the action of histamine and serotonin  
 CC (in e.g. inflammation and gastric acid secretion), the detection,  
 CC quantification and removal of histamine or serotonin (in animals, plants,  
 CC cell cultures, food materials, or humans) and in the treatment of various  
 CC diseases and allergies (e.g. type I hypersensitivity reactions,  
 CC urticaria, asthma, allergic rhinitis (hay fever), atopic dermatitis,  
 CC insect bites and food and drug allergies, abnormal blood pressure,  
 CC migraine, psychological disorders, respiratory disease, and coronary  
 CC heart disease). Histamine may also be used to regulate cellular growth  
 CC and tissue repair. The molecules may also be used as components of  
 CC vaccines directed against blood-sucking ectoparasites  
 XX  
 SQ Sequence 203 AA;  
 Query Match 31.4%; Score 353; DB 2; Length 203;  
 Best Local Similarity 36.3%; Pred. No. 7.8e-27;  
 Matches 74; Conservative 41; Mismatches 81; Indels 8; Gaps 6;  
 Qy 1 KMQQVLLTFVSAALATQAEITTSKAG--ENPLMAHEELIKGYODAKWSIDGVSATY 58  
 Db 1 KQALLIANGVIAANTAPQAPSSPRNEPLEKTTM-HSKELKGYODAKWSINQVSTTY 59  
 Qy 59 LAKTTYENDTSGWSQKCLQVOEIERKEEDYTVTSVTFPNASSPIKYNTVETVAKV 118  
 Db 60 FLRSTYNNDS-VWGNKFTCLSVTSKYESTFTVEYNTTYKNQSQ--QWVSMSENVTAVQ 116  
 Qy 119 QYGYKNIRNAIEYOVGGGLNITDPLTFDGLCDVFYVPNADGCELMVKSHYKHPDY 178  
 Db 117 BEGY-DVKNITQWTTENTKFNQDVIYFTDQCTDVLIFPYKEDGYELMVRSDLIQNTPTC 175  
 Qy 179 CTFVNVFCADKRTYDIFNEECV 202  
 Db 176 CQFIQFDL-VALGRTYNISTPDCV 198  
 RESULT 5  
 AAW37447  
 ID AAW37447 standard; protein; 190 AA.  
 XX  
 AC AAW37447;  
 XX  
 DT 08-JUN-1998 (first entry)  
 XX  
 DE Tick vasoactive amine binding protein 2 FS-HBP2.  
 XX  
 KW Female-specific vasoactive amine binding protein 1; FS-HCP1; histamine;  
 KW serotonin; assay; antihistamine; anti-inflammatory; insect bite;  
 KW snake bite; scorpion bite; dermatitis; vaccine; transgenic animal; tick.  
 XX  
 OS *Rhipicephalus appendiculatus*.  
 XX  
 PH Key Location/Qualifiers

FT Peptide 1.19  
 XX /label= Sig\_peptide  
 EN W09744451-A2.  
 PD 27-NOV-1997.  
 XX 19-MAY-1997; 97MO-GB001372.  
 PF 18-MAY-1996; 96GB-00010484.  
 PR 18-APR-1997; 97GB-00007844.  
 XX (OXFO-) OXFORD VACS LTD.  
 PA Paesen GC, Nuttall PA;  
 PI WPI: 1998-018506/02.  
 DR N-PSDB; AAV00228.  
 XX  
 FT New vasoactive amine binding proteins and related nucleic acid, vectors -  
 PT transformed cells and transgenic animals, used for assaying or removing  
 PT histamine and as antihistamine or anti-inflammatory agents.  
 XX  
 PS Example 2; Fig 2; 44pp; English.  
 XX  
 CC This protein comprises tick Rhipicephalus appendiculatus (Ra) novel  
 CC female-specific histamine binding protein 2 (FS-HBP2). Its amino acid  
 CC sequence was deduced from a cDNA clone (see AAV00228) obtained from a  
 CC salivary gland cDNA library. FS-HBP1 (see AAW37446) and male-specific  
 CC HSP1 (see AAW37448) and a related protein, D-REH6 (see AAW37449) from  
 CC Dermacentor reticulatus, were also identified. These novel vasoactive  
 CC amine binding proteins (VABPs) can be expressed in host cells using e.g.  
 CC a baculovirus expression system. They can be used: (i) to assay histamine  
 CC (or other VA such as serotonin) in body fluids or cell culture  
 CC supernatants, e.g. to monitor the effect of allergens; (ii) for binding  
 CC VA, e.g. to remove histamine from blood, food, cell cultures etc.; (iii)  
 CC as an antihistamine or anti-inflammatory agents, e.g. for treating  
 CC insect, snake or scorpion bites or dermatitis, or as a carrier for slow  
 CC release of histamine-related compounds; (iv) in vaccines to protect for  
 CC against metazoan parasites, especially in animals; (v) as reagents for  
 CC studying inflammation, involvement of VA in ulcer formation or the immune  
 CC response etc. VABPs provide a more sensitive assay for histamine than low  
 CC affinity antibodies currently used. They may also be more effective and  
 CC safer than conventional antihistamines  
 XX  
 SQ Sequence 190 AA;  
 Query Match 25.8%; Score 290.5; DB 2; Length 190;  
 Best Local Similarity 36.0%; Pred. No. 1.2e-20;  
 Matches 72; Conservative 32; Mismatches 79; Indels 17; Gaps 9;

7 LLLTFVSAALATQAEITTSKAGENPLMAHEELGKYQDAKMSIDQGVSVTVYIAKTYEN 66  
 DB 3 LLLTSLALVLA-----LSQVKNQPMADNANGAHQDAKMSLKADVENYVYMKATYKN 57  
 QY 67 DTGSMGSOFCLOVEIERKEDYTVTSVTF-RAASPIKRYNTETVKAFOYGYKNI 125  
 DB 58 DP-VWGNDFTCVGMADVNEDEKSIQAEFLFNANNADNMQF--ATEKTVAKMYGY-NR 113  
 QY 126 RNALEYOVGGGLNTDTLLFTDGLGDFVFPVNA--OGCELMVKSHYKHYGVDPYCTFV 182  
 DB 114 ENAFRYETEDGQVFTVIAVSD-DNCDVYIVGTGDNGBEYELMT--TDYDNIPIANCANK 170  
 QY 183 FNVFCADRKTYDIFNEECV 202  
 DB 171 FNEY-AVGRETRDVFTSACL 189

RESULT 6  
 AAY18079 standard; protein, 190 AA.  
 ID AAY18079  
 XX  
 AC AAY18079;

XX 06-AUG-1999 (first entry)  
 DT Histamine binding protein FS-HBP2.  
 XX  
 DE Histamine binding protein; serotonin binding compound; inflammation;  
 XX gastric acid secretion; allergy; type I hypersensitivity reaction;  
 KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;  
 KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;  
 KW respiratory disease; coronary heart disease; cellular growth regulator;  
 KW tissue repair; blood-sucking ectoparasite; therapy.  
 XX  
 OS Rhipicephalus appendiculatus.  
 XX  
 EN W09927104-A1.  
 PD 03-JUN-1999.  
 XX  
 PF 26-NOV-1998; 98MO-GB003530.  
 XX  
 PR 26-NOV-1997; 97GB-00025046.  
 PR 26-JUN-1998; 98GB-00013917.  
 XX  
 PA (OXFO-) OXFORD VACS LTD.  
 XX  
 PI Nuttall PA, Paesen GC;  
 XX  
 DR WPI: 1999-357841/30.  
 DR N-PSDB; AAX76965.  
 XX  
 PT Histamine and serotonin binding compounds useful for the treatment of  
 PT allergies.  
 XX  
 PS Claim 1; Fig 2; 84pp; English.  
 XX  
 CC This sequence is an example of a histamine or serotonin binding compound  
 CC (A), of the invention. The compounds are useful for regulating the action  
 CC of histamine and serotonin (in e.g. inflammation and gastric acid  
 CC secretion), the detection, quantification and removal of histamine or  
 CC serotonin (in animals, plants, cell cultures, food materials, or humans)  
 CC and in the treatment of various diseases and allergies (e.g. type I  
 CC hypersensitivity reactions, urticaria, asthma, allergic rhinitis (hay  
 CC fever), atopic dermatitis, insect bites and food and drug allergies,  
 CC abnormal blood pressure, migraine, psychological disorders, respiratory  
 CC disease, and coronary heart disease). Histamine may also be used to  
 CC regulate cellular growth and tissue repair. The molecules may also be  
 CC used as components of vaccines directed against blood-sucking  
 CC ectoparasites  
 XX  
 SQ Sequence 190 AA;  
 Query Match 25.8%; Score 290.5; DB 2; Length 190;  
 Best Local Similarity 36.0%; Pred. No. 1.2e-20;  
 Matches 72; Conservative 32; Mismatches 79; Indels 17; Gaps 9;

7 LLLTFVSAALATQAEITTSKAGENPLMAHEELGKYQDAKMSIDQGVSVTVYIAKTYEN 66  
 DB 3 LLLTSLALVLA-----LSQVKNQPMADNANGAHQDAKMSLKADVENYVYMKATYKN 57  
 QY 67 DTGSMGSOFCLOVEIERKEDYTVTSVTF-RAASPIKRYNTETVKAFOYGYKNI 125  
 DB 58 DP-VWGNDFTCVGMADVNEDEKSIQAEFLFNANNADNMQF--ATEKTVAKMYGY-NR 113  
 QY 126 RNALEYOVGGGLNTDTLLFTDGLGDFVFPVNA--OGCELMVKSHYKHYGVDPYCTFV 182  
 DB 114 ENAFRYETEDGQVFTVIAVSD-DNCDVYIVGTGDNGBEYELMT--TDYDNIPIANCANK 170  
 QY 183 FNVFCADRKTYDIFNEECV 202  
 DB 171 FNEY-AVGRETRDVFTSACL 189

RESULT 7





XX WO200115719-A2.  
 XX 08-MAR-2001.  
 XX 24-AUG-2000; 2000WO-GB003282.  
 XX 01-SEP-1999; 99GB-00020674.  
 XX (EVOL-) EVOLUTEC LTD.  
 XX Nuttall PA, Paesen GC;  
 XX WPI, 2001-257675/26.  
 XX Use of histacalin proteins for treating or preventing non-infective  
 PT conjunctivitis, or for manufacturing a medicament for treating or  
 PT preventing conjunctivitis, e.g. seasonal or perennial allergic  
 PT conjunctivitis.  
 XX Claim 4; Page 5-6; 1999; English.  
 XX The present invention relates to the use of a histacalin protein for  
 CC treating or preventing conjunctivitis. The present sequence is the  
 CC histacalin protein FS-HBP1. The invention is particularly useful in the  
 CC treatment of allergic or seasonal conjunctivitis  
 XX Sequence 171 AA;  
 SQ  
 Query Match 25.1%; Score 282; DB 4; Length 171;  
 Best Local Similarity 37.5%; Pred. No. 7.6e-20;  
 Matches 66; Conservative 29; Mismatches 69; Indels 12; Gaps 8;  
 QY 31 PLMAHELLGKYQAWMSIDGVSVTYLAKTETENDTSGMSQFCLQVQETEREEDY 90  
 DB 3 PDMADEANGAHQDMKSLKADVENYTMKATYKNDP-VWGNDFTCVGMANDVDEDK 61  
 QY 91 TVTSVTFP-RNASSPIKYYNTETVKAVFQYGYKNINALEYGVGGINTDTLTGSE 149  
 DB 62 SIOAEFLFMNNADTNMGP--ATEKVTAVKMGY-NRENARFETEDGQVFTDVIAYSD-D 117  
 QY 150 LCDVYFYVNAD---QGCETLWYKSHYGHVDPYCFVNVFCADRKTYDIFNECV 202  
 DB 118 NCDVITYPGTDGNEGEYELMT--TDYDNIPANCLINKNEY-AVGRETRDVFYSACL 170  
 RESULT 10  
 AAM37448  
 ID AAM37448 standard; protein; 200 AA.  
 XX AAM37448;  
 AC  
 XX 08-JUN-1998 (first entry)  
 DT  
 XX Tick vasoactive amine binding protein 1 MS-HBP1.  
 DE  
 XX Male-specific vasoactive amine binding protein 1; MS-HCP1; histamine;  
 KW serotonin; assay; antihistamine; anti-inflammatory; insect bite;  
 KW snake bite; scorpion bite; dermatitis; vaccine; transgenic animal; tick.  
 XX Rhipicephalus appendiculatus.  
 OS  
 XX Key Location/Qualifiers  
 FH 1..18  
 FT Peptide /label= Sig\_peptide  
 FT Modified-site /note= "Asn is N-glycosylated"  
 FT 79..81  
 XX WO9744451-A2.  
 XX 27-NOV-1997.  
 XX 19-MAY-1997; 97WO-GB001372.

XX 18-MAY-1996; 96GB-00010494.  
 XX 18-APR-1997; 97GB-00007844.  
 XX (OXFO-) OXFORD VACS LTD.  
 XX Paesen GC, Nuttall PA;  
 XX WPI, 1998-018506/02.  
 XX N-PSDB; AAV00229.  
 XX New vasoactive amine binding proteins and related nucleic acid, vectors -  
 PT transformed cells and transgenic animals, used for assaying or removing  
 PT histamine and as antihistamine or anti-inflammatory agents.  
 XX Example 2; Fig 3; 44pp; English.  
 XX This protein comprises tick Rhipicephalus appendiculatus (Ra) novel male-  
 CC specific histamine binding protein 1 (MS-HBP1). Its amino acid sequence  
 CC was deduced from a cDNA clone (see AAV00227) obtained from a salivary  
 CC gland cDNA library. Female-specific HSP1 and HSP2 (see AAM37446-47) and a  
 CC related protein, D.RHF6 (see AAM37449) from Dermacentor reticulatus, were  
 CC also identified. These novel vasoactive amine binding proteins (VABPs)  
 CC can be expressed in host cells using e.g. a baculovirus expression  
 CC system. They can be used: (i) to assay histamine (or other VA such as  
 CC serotonin) in body fluids or cell culture supernatants, e.g. to monitor  
 CC the effect of allergens; (ii) for binding VA, e.g. to remove histamine  
 CC from blood, food, cell cultures etc.; (iii) as an antihistamine or anti-  
 CC inflammatory agents, e.g. for treating insect, snake or scorpion bites or  
 CC dermatitis or as a carrier for slow release of histamine-related  
 CC compounds; (iv) in vaccines to protect against metazoan parasites,  
 CC especially in animals; (v) as reagents for studying inflammation,  
 CC involvement of VA in ulcer formation or the immune response etc. VABPs  
 CC provide a more sensitive assay for histamine than low-affinity antibodies  
 CC currently used. They may also be more effective and safer than  
 CC conventional antihistamines  
 XX Sequence 200 AA;  
 SQ  
 Query Match 23.6%; Score 265; DB 2; Length 200;  
 Best Local Similarity 34.5%; Pred. No. 4.7e-18;  
 Matches 69; Conservative 34; Mismatches 71; Indels 26; Gaps 10;  
 QY 3 MOVVLLTFYSALATQAEFTSAKAGENPLMAHELLGKYQAWMSIDGVSVTYLAKT 62  
 DB 1 MKVLLLV--IGALLCONADA-----NTWANEALGSGYQAWMSIQDQKRYTLAQA 51  
 QY 63 TYENDTSGMSQFCLQV--QETEREEDYTVTSVTFPNAS-SPIKYNTVETVKAVFQ 119  
 DB 52 TQTTD-GVWGEFFTCVSVTAETKIGKKLNATI---LYNKHLLTDLKESHEITTYWKAD 106  
 QY 120 YGYKNINALEYGVGGINT--DTLIFTDGEICDVFYVFP---NADQG-CELVVKKSHYK 173  
 DB 107 Y--TTENGIKETIGRTGTQTFEDVVFSDYKNCVYIFPKRGSGEGYELWVSRDXID 163  
 QY 174 HVPDYCTFVNVFCADRKXT 193  
 DB 164 KIPDCKFTWAVFAQOQCKT 183  
 RESULT 11  
 AAY18080  
 ID AAY18080 standard; protein; 200 AA.  
 XX AAY18080;  
 AC  
 XX 06-AUG-1999 (first entry)  
 DT  
 XX Histamine binding protein MS-HBP1.  
 DE  
 XX Histamine binding protein; serotonin binding compound; inflammation;  
 KW gastric acid secretion; allergy; type I hypersensitivity reaction;  
 KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;

KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;  
 KM respiratory disease; coronary heart disease; cellular growth regulator;  
 KW tissue repair; blood-sucking ectoparasite; therapy.  
 XX  
 OS Rhipicephalus appendiculatus.  
 XX  
 PN WO927104-A1.  
 XX  
 PD 03-JUN-1999.  
 XX  
 PF 26-NOV-1996; 98WO-GB003530.  
 XX  
 PR 26-NOV-1997; 97GB-00025046.  
 XX  
 PR 26-JUN-1998; 98GB-00013917.  
 XX  
 PA (OXFO-) OXFORD VACS LTD.  
 XX  
 PI Nuttall PA, Paesen GC;  
 XX  
 PI WPI; 1999-357841/30.  
 DR N-PSDB; AAX76966.  
 DR  
 PT Histamine and serotonin binding compounds useful for the treatment of  
 XX allergies.  
 PS  
 PS Claim 1; Fig 3; 84pp; English.  
 XX  
 XX This sequence is an example of a histamine or serotonin binding compound  
 CC (A), of the invention. The compounds are useful for regulating the action  
 CC of histamine and serotonin (in e.g. inflammation and gastric acid  
 CC secretion), the detection, quantification and removal of histamine or  
 CC serotonin (in animals, plants, cell cultures, food materials, or humans)  
 CC and in the treatment of various diseases and allergies (e.g. type I  
 CC hypersensitivity reactions, urticaria, asthma, allergic rhinitis (hay  
 CC fever), atopic dermatitis, insect bites and food and drug allergies,  
 CC abnormal blood pressure, migraine, psychological disorders, respiratory  
 CC disease, and coronary heart disease). Histamine may also be used to  
 CC regulate cellular growth and tissue repair. The molecules may also be  
 CC used as components of vaccines directed against blood-sucking  
 CC ectoparasites  
 CC  
 SQ Sequence 200 AA;  
 Query Match 23.6%; Score 265; DB 2; Length 200;  
 Best Local Similarity 34.5%; Pred. No. 4.7e-18;  
 Matches 69; Conservative 34; Mismatches 71; Indels 26; Gaps 10;  
 QY 3 MOVVLLITFVSALALQVARTTSKAGENPLMAHELLKCYODAMKSIDQGVSVTVLAKT 62  
 DB 1 MKVLLIV-LGALCONADA-----NPTWANEAKLGSYQAMSLQDDQKRYLLQA 51  
 QY 63 TYENDTSGWSGPFKQV--QEIERKEEDYTVSVTFPNAS-SPIKYNTVETVKAVFQ 119  
 DB 52 TQTTD-GVWGEFFTCVSTAEKIGKKLNATI---LYKKEHLTLKESHTITVWKAID 106  
 QY 120 YGVKNINALEYGVGGSLNT--DTLITFDGELCDVYFVR--NMDQG-CEIMVKKSHYK 173  
 DB 107 Y--TTENGIRKYETQGRRTQTFEDVFSDYKNCVIFVPERGSDGDEGLWVSEDKID 163  
 QY 174 HVPDYCTFVFNVCADKDKT 193  
 DB 164 KIPDCKFTMAVFAQOQEK 183  
 RESULT 12  
 AAM37446  
 ID AAM37446 standard; protein; 190 AA.  
 AC AAM37446;  
 XX  
 DT 08-JUN-1998 (first entry)  
 XX  
 DE Tick vasoactive amine binding protein 1 FS-HBP1.

XX Female-specific vasoactive amine binding protein 1; FS-HBP1; histamine;  
 KM serotonin; assay; antihistamine; anti-inflammatory; insect bite;  
 KW snake bite; scorpion bite; dermatitis; vaccine; transgenic animal; tick.  
 XX  
 OS Rhipicephalus appendiculatus.  
 XX  
 FH Key  
 FT Peptide  
 FT 1-18  
 FT /label= Sig\_peptide  
 XX  
 PN WO9744451-A2.  
 XX  
 PD 27-NOV-1997.  
 XX  
 PF 19-MAY-1997; 97WO-GB001372.  
 XX  
 PR 18-MAY-1996; 96GB-00010484.  
 XX  
 PR 18-APR-1997; 97GB-00007844.  
 XX  
 PA (OXFO-) OXFORD VACS LTD.  
 XX  
 PI Paesen GC, Nuttall PA;  
 XX  
 PI WPI; 1998-018506/02.  
 DR N-PSDB; AAV00227.  
 DR  
 PT New vasoactive amine binding proteins and related nucleic acid, vectors  
 PT transformed cells and transgenic animals, used for assaying or removing  
 PT histamine and as antihistamine or anti-inflammatory agents.  
 XX  
 XX Example 2; Fig 1; 44pp; English.  
 XX  
 XX This protein comprises tick Rhipicephalus appendiculatus (Ra) novel  
 CC female-specific histamine binding protein 1 (FS-HBP1). Its amino acid  
 CC sequence was deduced from a cDNA clone (see AAV00227) obtained from a  
 CC salivary gland cDNA library. FS-HBP1 and male-specific HSP1 (see AAM37447  
 CC -48) and a related protein, D.REM (see AAM37449) from Dermacentor  
 CC reticulatus, were also identified. These novel vasoactive amine binding  
 CC proteins (VABPs) can be expressed in host cells using e.g. a baculovirus  
 CC expression system. They can be used: (i) to assay histamine (or other VA  
 CC such as serotonin) in body fluids or cell culture supernatants, e.g. to  
 CC monitor the effect of allergens; (ii) for binding VA, e.g. to remove  
 CC histamine from blood, food, cell cultures etc.; (iii) as an antihistamine  
 CC or anti-inflammatory agent, e.g. for treating insect, snake or scorpion  
 CC bites or dermatitis, or as a carrier for slow release of histamine-  
 CC related compounds; (iv) in vaccines to protect against metazoan  
 CC parasites, especially in animals; (v) as reagents for studying  
 CC inflammation, involvement of VA in ulcer formation or the immune response  
 CC etc. VABPs provide a more sensitive assay for histamine than low-affinity  
 CC antibodies currently used. They may also be more effective and safer than  
 CC conventional antihistamines  
 CC  
 SQ Sequence 190 AA;  
 Query Match 22.4%; Score 252; DB 2; Length 190;  
 Best Local Similarity 33.5%; Pred. No. 8.7e-17;  
 Matches 68; Conservative 28; Mismatches 89; Indels 18; Gaps 8;  
 QY 3 MOVVLLITFVSALALQVARTTSKAGENPLMAHELLKCYODAMKSIDQGVSVTVLAKT 62  
 DB 1 MKLLISIAFVTLASQVKA-----DKVWADKANGSHQAMHLLQKLYBENDYLKA 52  
 QY 63 TYENDTSGWSGPFKQVQEIIRKEEDYTVSVTFPNAS-SPIKYNTVETVKAVFQY 122  
 DB 53 TYKNDP-VGNDPTCTGTAQNLNDEKKNVZAMFPMNADTV-YQHFEEKATPDKNVGY 110  
 QY 123 KNIRNAIEYGVGGSLNTITTLITFDGELCDVYFVNAD--QGCELMVKKSHYKVPYC 179  
 DB 111 -NKENAITVQTEDEGQVLTDLVAFSD-DNCYVYVLAGPDGSGAGYELMA--IDYTVFASC 166  
 QY 180 TFEVNFCAKDKRTYDIFNEBCV 202



XX 20-JUN-2001 (first entry)  
DT Histacalin protein MS-HBP1.  
XX  
DE Histacalin; MS-HBP1, conjunctivitis.  
XX  
XX Unidentified.  
OS  
XX WO200115719-A2.  
PN  
XX 08-MAR-2001.  
PD  
XX 24-AUG-2000; 2000WO-GB003282.  
PF  
XX 01-SEP-1999; 99GB-00020674.  
PR  
XX (EVOL-) EVOLUTEC LTD.  
PA  
XX Nuttall PA, Paesen GC;  
PI  
XX WPI; 2001-257675/26.  
DR  
XX  
XX Use of histacalin proteins for treating or preventing non-infective  
PT conjunctivitis, or for manufacturing a medicament for treating or  
PT preventing conjunctivitis, e.g. seasonal or perennial allergic  
PT conjunctivitis.  
XX  
XX Claim 4; Page 5-6; 19pp; English.  
PS  
XX  
XX The present invention relates to the use of a histacalin protein for  
CC treating or preventing conjunctivitis. The present sequence is the  
CC histacalin protein FS-HBP1. The invention is particularly useful in the  
CC treatment of allergic or seasonal conjunctivitis  
XX  
SQ Sequence 182 AA;

Query Match 22.4%; Score 251.5; DB 4; Length 182;  
Best Local Similarity 35.3%; Pred. No. 9.2e-17;  
Matches 61; Conservative 29; Mismatches 66; Indels 17; Gaps 8;  
QY 30 NPLMAHEELGKIYQAMKSIDQGVSVTVYLAITYENDTSGMSQFKLOY--QELRKE 87  
DB 1 NPTMANEAKDGSYQAMKSLQDDCKRYTLAQTQTD-GYWGSEFTCVSTAKIGKKK 59  
QY 88 EDYTVSVTFENAS-SPIKYNVTEYKAVFOYGVKIRNAIEYVGGLNT--DTLI 144  
DB 60 LNATL---LYNKHLLTDKESHERITVWKAYDY---TTENGKITYGTGTRTQTFEDVYV 112  
QY 145 FTDSGLCDVYVP--NADQG-CELVKSKSHYKVPDYCTFVFNVPQAKDKRT 193  
DB 113 FSDYKNCDIVIFVPERGSGDEGDEYELWVSEDKIDKIPDCKFTWYVFAQOOEKT 165

Search completed: March 9, 2004, 13:33:28  
Job time : 62 secs